STEVENS' HANDBOOK OF EXPERIMENTAL PSYCHOLOGY AND COGNITIVE NEUROSCIENCE

FOURTH EDITION

VOLUME 4: Developmental & Social Psychology

Simona Ghetti
VOLUME EDITOR

John T. Wixted
EDITOR-IN-CHIEF

STEVENS' HANDBOOK OF EXPERIMENTAL PSYCHOLOGY AND COGNITIVE NEUROSCIENCE

STEVENS' HANDBOOK OF EXPERIMENTAL PSYCHOLOGY AND COGNITIVE NEUROSCIENCE

FOURTH EDITION

Volume 4
Developmental & Social Psychology

Editor-in-Chief

JOHN T. WIXTED

Volume Editor

SIMONA GHETTI

WILEY

This book is printed on acid-free paper. ∞

Published by John Wiley & Sons, Inc.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data

The Library of Congress has cataloged the combined volume as follows:

Name: Wixted, John T., editor.
Title: Stevens' handbook of experimental psychology and cognitive neuroscience / by John T. Wixted (Editor-in-chief).
Other titles: Handbook of experimental psychology.
Description: Fourth edition. | New York : John Wiley & Sons, Inc., [2018] | Includes index. Contents: Volume 1. Learning and memory – Volume 2. Sensation, perception, and attention – Volume 3. Language & thought – Volume 4. Developmental & social psychology – Volume 5. Methodology.
Identifiers: LCCN 2017032691 | ISBN 9781119170013 (cloth : vol. 1) |
ISBN 9781119170037 (epdf : vol. 1) | ISBN 9781119170020 (epub : vol. 1) |
ISBN 9781119170044 (cloth : vol. 2) | ISBN 9781119174158 (epdf : vol. 2) |
ISBN 9781119174073 (epub : vol. 2) | ISBN 9781119170693 (cloth : vol. 3) |
ISBN 9781119170730 (epdf : vol. 3) | ISBN 9781119170716 (epub : vol. 3) |
ISBN 9781119170051 (cloth : vol. 4) | ISBN 9781119170068 (epdf : vol. 4) |
ISBN 9781119170082 (epub : vol. 4) | ISBN 9781119170129 (cloth : vol. 5) |
ISBN 9781119170150 (epdf : vol. 5) | ISBN 9781119170143 (epub : vol. 5)
Subjects: LCSH: Psychology, Experimental. | Cognitive neuroscience.
Classification: LCC BF181 .H336 2018 | DDC 150—dc23 LC record available at https://lccn.loc.gov/2017032691

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books. For more information about Wiley products, visit our web site at www.wiley.com.

Printed in the United States of America.
10 9 8 7 6 5 4 3 2 1

Contributors

Dima Amso
Brown University

Jennifer S. Beer
University of Texas at Austin

Steven T. Bengal
Ohio State University

Lindsay C. Bowman
University of California, Davis

Laura Martin Braunstein
Columbia University

Jeffrey A. Brooks
New York University

Christine Coughlin
University of California, Davis

Eveline A. Crone
Universiteit Leiden Faculteit Sociale Wetenschappen, Leiden, Zuid-Holland

Audun Dahl
University of California, Santa Cruz

Mauricio R. Delgado
Rutgers University Newark College of Arts and Sciences

Jonathan B. Freeman
New York University

Amber M. Gaffney
Humboldt State University

Bertram Gawronski
University of Texas at Austin

Simona Ghetti
University of California, Davis

Jeremy D. Gretton
Ohio State University

Adam Hahn
University of Cologne, Koeln, Germany

Cindy Harmon-Jones
The University of New South Wales

Eddie Harmon-Jones
The University of New South Wales

Paul Hastings
University of California, Davis

Michael A. Hogg
Claremont Graduate University

Teresa Iuculano
Stanford University

Melanie Killen
University of Maryland

David Klahr
Carnegie Mellon University

Jessica E. Koski
University of Texas at Austin

Sarah Leckey
University of California, Davis

Jeffrey Lidz
University of Maryland at College Park

Vinod Menon
Stanford University

Lisa Oakes
University of California, Davis

Kevin N. Ochsner
Columbia University

Koraly Pérez-Edgar
Pennsylvania State University

Laurel Perkins
University of Maryland at College Park

Tom F. Price
Army Research Laboratory, Adelphi, Maryland

Anastasia E. Rigney
University of Texas at Austin

Mark A. Sabbagh
Queen's University

Vladimir M. Sloutsky
Ohio State University

Megan E. Speer
Rutgers Biomedical and Health Sciences

Duane T. Wegener
Ohio State University

Kiki Zanolie
Institute of Psychology, Leiden University, Leiden, The Netherlands

Corinne Zimmerman
Illinois State University

Contents

Preface

Since the first edition was published in 1951, *The Stevens' Handbook of Experimental Psychology* has been recognized as the standard reference in the experimental psychology field. The most recent (third) edition of the handbook was published in 2004, and it was a success by any measure. But the field of experimental psychology has changed in dramatic ways since then. Throughout the first three editions of the handbook, the changes in the field were mainly quantitative in nature. That is, the size and scope of the field grew steadily from 1951 to 2004, a trend that was reflected in the growing size of the handbook itself: the one-volume first edition (1951) was succeeded by a two-volume second edition (1988) and then by a four-volume third edition (2004). Since 2004, however, this still-growing field has also changed qualitatively in the sense that, in virtually every subdomain of experimental psychology, theories of the mind have evolved to include theories of the brain. Research methods in experimental psychology have changed accordingly and now include not only venerable EEG recordings (long a staple of research in psycholinguistics) but also MEG, fMRI, TMS, and single-unit recording. The trend toward neuroscience is an absolutely dramatic, worldwide phenomenon that is unlikely ever to be reversed. Thus, the era of purely behavioral experimental psychology is already long gone, even though not everyone has noticed. Experimental psychology and *cognitive neuroscience* (an umbrella term that, as used here, includes behavioral neuroscience, social neuroscience, and developmental neuroscience) are now inextricably intertwined. Nearly every major psychology department in the country has added cognitive neuroscientists to its ranks in recent years, and that trend is still growing. A viable handbook of experimental psychology should reflect the new reality on the ground.

There is no handbook in existence today that combines basic experimental psychology and cognitive neuroscience, despite the fact that the two fields are interrelated—and even interdependent—because they are concerned with the same issues (e.g., memory, perception, language, development, etc.). Almost all neuroscience-oriented research takes as its starting point what has been learned using behavioral methods in experimental psychology. In addition, nowadays, psychological theories increasingly take into account what has been learned about the brain (e.g., psychological models increasingly need to be neurologically plausible). These considerations explain why I chose a new title for the handbook: *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience.* This title serves as a reminder that the two fields go together and as an announcement that the *Stevens' Handbook* now covers it all.

The fourth edition of the *Stevens' Handbook* is a five-volume set structured as follows:

1. **Learning & Memory:** Elizabeth A. Phelps and Lila Davachi (volume editors)

   Topics include fear learning, time perception, working memory, visual object recognition, memory and future imagining, sleep and memory, emotion and memory, attention and memory, motivation and memory, inhibition in memory, education and memory, aging and memory, autobiographical memory, eyewitness memory, and category learning.
2. **Sensation, Perception, & Attention:** John T. Serences (volume editor)

   Topics include attention; vision; color vision; visual search; depth perception; taste; touch; olfaction; motor control; perceptual learning; audition; music perception; multisensory integration; vestibular, proprioceptive, and haptic contributions to spatial orientation; motion perception; perceptual rhythms; the interface theory of perception; perceptual organization; perception and interactive technology; and perception for action.
3. **Language & Thought:** Sharon L. Thompson-Schill (volume editor)

   Topics include reading, discourse and dialogue, speech production, sentence processing, bilingualism, concepts and categorization, culture and cognition, embodied cognition, creativity, reasoning, speech perception, spatial cognition, word processing, semantic memory, and moral reasoning.
4. **Developmental & Social Psychology:** Simona Ghetti (volume editor)

   Topics include development of visual attention, self-evaluation, moral development, emotion-cognition interactions, person perception, memory, implicit social cognition, motivation group processes, development of scientific thinking, language acquisition, category and conceptual development, development of mathematical reasoning, emotion regulation, emotional development, development of theory of mind, attitudes, and executive function.
5. **Methodology:** Eric-Jan Wagenmakers (volume editor)

   Topics include hypothesis testing and statistical inference, model comparison in psychology, mathematical modeling in cognition and cognitive neuroscience, methods and models in categorization, serial versus parallel processing, theories for discriminating signal from noise, Bayesian cognitive modeling, response time modeling, neural networks and neurocomputational modeling, methods in psychophysics analyzing neural time series data, convergent methods of memory research, models and methods for reinforcement learning, cultural consensus theory, network models for clinical psychology, the stop-signal paradigm, fMRI, neural recordings, and open science.

How the field of experimental psychology will evolve in the years to come is anyone's guess, but the *Stevens' Handbook* provides a comprehensive overview of where it stands today. For anyone in search of interesting and important topics to pursue in future research, this is the place to start. After all, you have to figure out the direction in which the river of knowledge is currently flowing to have any hope of ever changing it.

DEVELOPMENTAL PSYCHOLOGY

CHAPTER 1

Development of Visual Attention

LISA OAKES AND DIMA AMSO

INTRODUCTION

Consider a child searching a crowded room for her parent. Perhaps there are several people in the room as well as furniture, toys, and other objects. In addition, there may be decorations on the wall, light fixtures hanging from the ceiling, windows, curtains, and so on. *Visual attention* is the set of processes that allows the child to filter the overly cluttered visual world, selecting some available information to process—in this case the people—and inhibiting other available information—in this case the furniture, light fixtures, and curtains. These attentional processes are governed by a complex set of interacting neural systems that develop over infancy and childhood.

In what follows, we provide formal definitions of those visual attention processes that are most relevant to infants and children. Next, we describe influential models and tasks of visual attention. Then we discuss what is known about the development of attentional processes during infancy, early childhood, and later childhood and beyond. We describe historical work examining looking behavior as a measure of visual attention, which provides a foundation for our understanding of the development of visual attention across childhood. We also discuss more contemporary work using more standard visual attention tasks, often adapted from work with adults. Throughout, we discuss the paradigms that have been used to assess visual attention in infancy and childhood, including a discussion of what specific computations or processes of visual attention each assesses. Finally, we examine how visual attention processes (and their development) interact with other cognitive and perceptual systems such as memory and learning, how novel neuroimaging tools add insight into neural systems development underlying visual attention, and future directions in visual attention research.

BACKGROUND ISSUES

Defining Visual Attention

Defining attention is not trivial. In part, this is because many meanings of the term “attention” are intuitive—we know that children who are paying attention are quiet, looking at the thing they are paying attention to, and not doing something else. We know that children who have problems with attention have difficulty staying on task and are easily distracted by thoughts, tasks, or stimuli in their environment. We command others to “pay attention,” and we talk informally about the inability to maintain attention (e.g., “spacing out”).

However, the scientific study of the development of attention requires a more formal and precise definition. As the example just described illustrates, attention is necessary in contexts of information overload. Without attention, it would be impossible to bind features of visual objects (such as color and shape) (Treisman, 1998), overcome limited visual working memory capacity (Awh, Vogel, & Oh, 2006), or process a signal effectively in a noisy context (Carrasco, 2014). Luck and Vecera (2002) offer a process-oriented definition of attention that states that (1) attention is the selection of information among alternatives, and (2) this selection improves the effectiveness of mental processes. Visual attention, therefore, allows us to *select* information from the visual environment for further processing while simultaneously ignoring or inhibiting competing information that is not selected. The point is that when defining the term "attention," we can focus on the function of attention. By engaging in selection and inhibition, visual attention turns up the gain on some items and locations for subsequent goal-relevant action, perception, and memory (Carrasco, 2011, 2014; Markant, Worden, & Amso, 2015; Zhang et al., 2011).

Note, however, that this definition of attention does not restrict attention to a single modality or level of processing. Our task here, however, is to describe the development of *visual* attention. It is important to recognize that even behavior that we would clearly consider visual attention—for example, directing fixation or processing resources to an aspect of the visual environment—is a function of many processes, only some of which are solely visual. General level of arousal, for example, may influence the depth of one's attentional engagement. Voluntary control over head and eye movements will contribute to overt direction of visual attention. And high-level processes, such as establishing goals, prioritizing events and stimuli in terms of their relevance, and applying existing knowledge to a current situation, will influence visual attention. As such, visual attention does not operate in isolation. Recognizing these connections and evaluating the literature with an understanding of the possible roles of multiple factors and processes on visual attention can enable us to attain deep understanding of visual attention and its development.

It is also important to recognize that visual attention is a set of computations or processes rather than a skill or content domain. A formal and precise definition of attention requires consideration of the structures and mechanisms that support these processes and functions. An important framework for understanding visual attention is Posner and Petersen's (Petersen & Posner, 2012; Posner & Petersen, 1990) classic model. This model describes three aspects of attention—alerting, orienting, and executive attention—that are supported by different neural networks (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005; Fan, McCandliss, Sommer, Raz, & Posner, 2002; Posner & Petersen, 1990). Each of these aspects of attention applies to specific aspects of visual attention. The alerting response, supported by thalamic involvement, is a phasic attentional readiness and is a prepared response to a warning (a tone prepares runners for the official start of a race) stimulus. A related sustained attention mechanism involves a more continuous focus on a particular task or stimulus. The orienting mechanism involves shifting attention to an item or a location either with an overt eye movement or covertly, without a physical eye movement. Visual attention orienting recruits a parietal network. The executive attention mechanism is involved in switching, inhibiting, and general top-down control of visual attention, and it involves frontoparietal cortices and the anterior cingulate cortex. Clearly, each of these attention functions also is influenced by and relies on other processes.

For example, motor development and oculomotor development are extremely relevant to the development of visual attention processes. Overt attention, which in some ways is the most straightforward and obvious example of visual attention, involves turning one's head and eyes to bring a stimulus, object, or feature of the environment into focus. Overt attention thus relies on the physical abilities involved in holding one's head upright, making effortful and voluntary head turns, and voluntarily controlling eye movements. Motor control over the head and eyes undergoes significant developmental change in infancy (Bertenthal & Von Hofsten, 1998; Canfield & Kirkham, 2001; von Hofsten, 2004), which opens up novel exploratory and attentional strategies for young infants (Gibson, 1988).

Moreover, there are many similarities between visual attention and related general attention processes as well as attention that operates over other sensory modalities, such as auditory attention. For example, regardless of the modality, attention involves selection of relevant stimuli and inhibition of distractors. In addition, attention as used in one modality may in fact influence attention in other modalities. Amso et al. (2014) argued that the development of visual attention may depend on the development of visual processing (see also Amso & Scerif, 2015). Smith and Trainor (2011) made a similar argument with respect to auditory selective attention: specifically that auditory selective attention in infants depends on infants' ability to perceptually process target and nontarget sounds. Direct data comparing the developmental trajectories of these processes is sparse. One recent study (Günther et al., 2014) compared visual and auditory selective attention processes in a group of participants 7 to 77 years on a focused-attention task. The authors found that participants were better in the visual than in the auditory conditions, but the modality effect diminished with age. These data suggest different developmental trajectories for visual and auditory attention. We highlight these similarities and differences to point out that although understanding visual attention is relevant to the study of auditory attention, the two processes have distinct and nontransferable developmental trajectories.

Influential Models and Common Tasks

Most views of attention derive from the influential model of Posner and Petersen (Petersen & Posner, 2012; Posner & Petersen, 1990). As described in the previous section, this model describes alerting, orienting, and executive attention, all subserved by different neural structures and all of which have different functions related to the selection and filtering of relevant information and the inhibition of irrelevant or distracting information. These attentional processes have been widely studied and have been examined over a wide age range. Thus, many other models of attention have focused on similar processes.

As an example, consider the four functions of attention Colombo (2001) described in infancy. These four functions are closely related to Posner and Petersen's attention networks (Petersen & Posner, 2012; Posner & Petersen, 1990). Specifically, Colombo describes alertness, spatial orienting, attention to object features, and endogenous control. Here, the term "alertness" refers to Posner and Petersen's alerting network. It reflects the ability to both attain as well as maintain an alert state. The terms "spatial orienting" and "attention to object features" correspond to Posner and Petersen's orienting mechanism. Colombo separated this network into two functions—one for selecting and shifting attention to particular locations (spatial orienting) and another for selecting and shifting attention to particular types of objects features (perhaps their shape or color). This differentiation roughly corresponds to the "what" and "where" visual systems (Ungerleider & Pessosa, 2008). Finally, Colombo (2001) described endogenous attention, which corresponds to Posner and Petersen's executive attention. For Colombo, this is the ability to voluntarily direct attention to particular features or aspects of the environment as well as the ability to inhibit attending to some features or aspects of the environment. These functions correspond to top-down control over the other visual attention functions. Therefore, Colombo's model is specifically directed at explaining attention in infancy, but the components and functions of attention are clearly closely tied to the classic Posner and Petersen conception of attention networks.

The tasks commonly used to assess visual attention are designed to index the visual attention processes and networks described in the Posner and Petersen model (Petersen & Posner, 2012; Posner & Petersen, 1990). A standard procedure used to study visual attention across populations is the spatial cuing procedure (Posner, 1980). In this general class of tasks, attention processes are invoked with a cue. The cue may indicate that a target is about to occur, or it may indicate a potential location of the impending target. For example, Posner and colleagues developed the Attention Network Test (ANT) (Fan et al., 2002), which includes several types of trials that use cuing to access alerting, orienting, and executive attention networks. Participants are instructed to respond to an identified target item. To assess alerting, a cue warns the participant to prepare for the coming target but gives no information about the location that the target will occur (e.g., in Figure 1.1a, there are asterisks—or cues—in

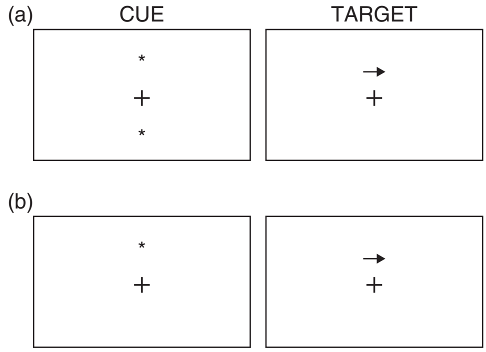


Figure 1.1 A schematic depiction of the Attention Network Task (ANT) (e.g., Fan et al., 2002). In each figure, the cross represents the fixation point, the asterisk is a cue, and the arrow is the target. The figures in (a) illustrate an alerting trial in which the asterisks act as a cue and alert the participant to prepare to respond to a target stimulus but provide no information to the location of that target. The figures in (b) illustrate a valid trial in which the cue indicates both that a target stimulus will occur and also the location in which it will occur, offering the participant the opportunity to covertly orient to that location and prepare a response.

both possible target locations). Thus, the presence of the cue invokes a phasic alerting response in preparation for the target stimulus but does not provide any useful information about how to selectively direct or control attention. To assess orienting, the cue also contains information about the location where the target stimulus will occur (e.g., in Figure 1.1b, there is only a single asterisk in the location where the target will later appear). This type of cue allows the participant to prepare for a target in a specific location, perhaps "covertly," or without an eye movement, shifting attention to the cued location in anticipation of the emergence of the target at that location.

Cuing is not the only way in which researchers have examined orienting attention. A common task used to understanding orienting is visual search (e.g., Treisman & Gelade, 1980). In such tasks, a target item is cast in the midst of varying numbers of

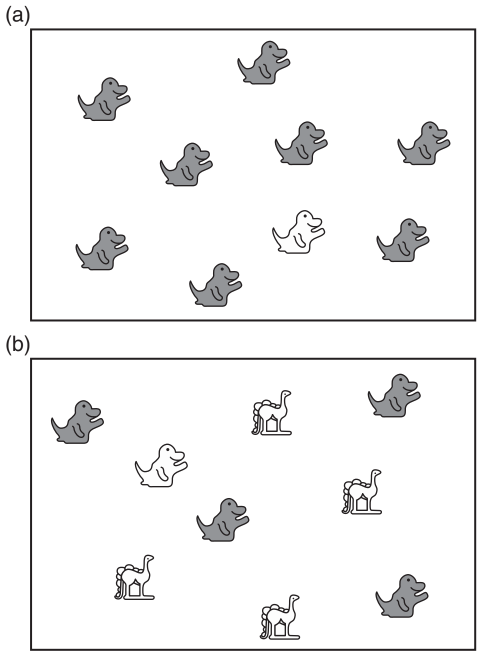


Figure 1.2 Examples of visual search arrays. In (a) the target is defined by a single feature (color), whereas in (b) the target is defined by the combination of two features (color and shape).
SOURCE: Reprinted from Gerhardstein & Rovee-Collier (2002). Copyright (2002) with permission from Elsevier.

distractors. If the target and distractor vary along only one feature dimension, as in Figure 1.2a, the target pops out and is considered preattentive (e.g., Treisman & Gelade, 1980); that is, the target can be detected and located even without the use of attention. One key characteristic of pop-out search is that increasing the number of distractors in the display does not result in longer search times to the target. When the target and distractors share a conjunction of features (Figure 1.2b), in contrast, visual search is effortful and requires attention. In this case, target identification is made progressively more effortful, as indexed by increasing target search times, by increasing the similarity (or competition) between the distractors and the target, or by increasing the number of distractors in the scene (e.g., Treisman & Gelade, 1980), suggesting that participants take longer to detect the target when they have to shift their attention to larger numbers of items. Variants of visual search have become widely used to understand attentional processes in infants (Adler, 2005), toddlers (Gerhardstein & Rovee-Collier, 2002; Scerif, Cornish, Wilding, Driver, & Karmiloff-Smith, 2004), and children (Donnelly et al., 2007). Indeed, some work has explored changes in attention across the life span by examining performance in visual search over a wide age range (Trick & Enns, 1998).

Assessment of executive attention requires that some perceptual conflict be resolved, and such tasks engage midline frontal areas and the lateral prefrontal cortex (Fan et al., 2002). In the ANT, for example, executive attention is assessed using a version of the Eriksen Flanker task (Eriksen & Eriksen, 1974). In this task, a target is an arrow presented in the center of a display. In the simple version of this task, the subject simply has to determine whether the arrow points to the right or the left. However, to assess executive attention, trials are presented in which the central arrow is “flanked” by distracting arrows. Figure 1.3 illustrates child-friendly versions of this task. In the “Fish” adaptation, for example, the trials presented on the left do not require executive attention because all the fish point in the same direction and thus no conflict needs to be resolved. On the trials presented on the left of the figure, in contrast, the flanking fish point one direction and the central fish points in the opposite direction. In this case, the central target and the flanker are conflicting. Because the child’s task is to report the direction the central fish (or arrow) is pointing, accurately responding in the flanker tasks requires inhibiting responding to the flanker fish (arrows) and focusing attention

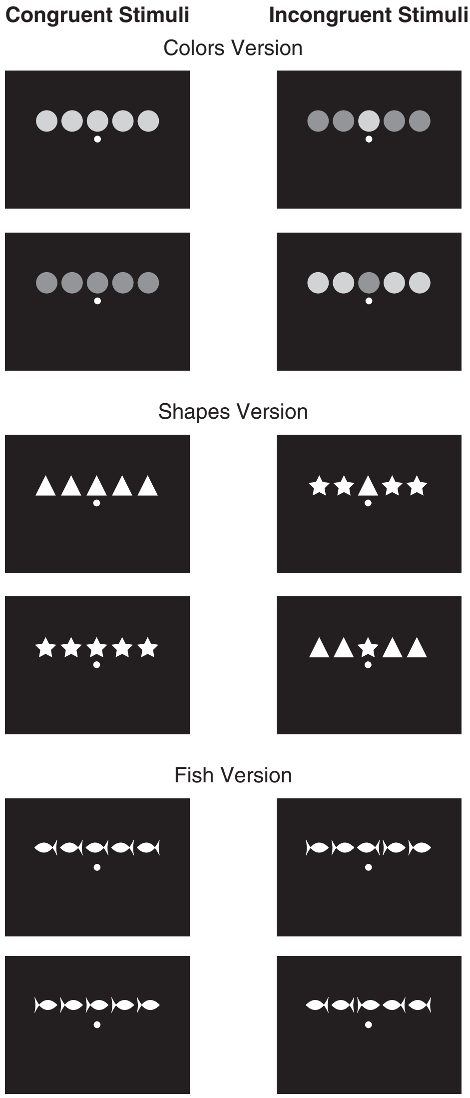


Figure 1.3 Examples of flanker tasks for children.
Source: Reprinted from McDermott, Perez-Edger, & Fox 2007. Copyright 2007 Psychonomic Society, Inc., with permission of Springer.

on the central fish (arrow). Fan et al. (2005) confirmed that the executive attention portion of the ANT engage different brain regions from the other portions of the ANT and that this flanker task engages frontoparietal and anterior cingulate regions generally thought to be involved when dealing with conflict. To better assess young children's performance on this task, McDermott, Perez-Edgar, and Fox (2007) used the variations presented in Figure 1.3 (see also Rueda et al., 2004) and demonstrated behavioral effects of the flankers on the performance of children between 4 and 6 years of age.

In sum, there is a large body of research presenting tasks to assess the development of attention. These tasks have been strongly influenced by the traditional model of attentional networks, originally proposed by Posner and Petersen (Petersen & Posner, 2012; Posner & Petersen, 1990). These visual attention tasks have proven to be powerful for studying visual attention beginning in infancy and extending to adulthood, as described next.

Development of Attention

Attention in Infancy

Different visual attention processes emerge beginning in infancy. However, our description of the ANT task and spatial cuing more generally should make it clear that many aspects or processes of attention are extremely difficult to measure and study in infancy. As a result, historically, the study of attention in infancy conflated attentional processes with measures used to index them, including looking times and oculomotor control, making the early study of visual attention in infancy actually the study of visual behavior in infancy. Indeed, a large number of studies were published in the 1960s and 1970s examining models of infant attention, the effect of stimulus properties on infant attention, and the relation between infant attention and memory.

In the first postnatal weeks, infants have difficulty initiating and maintaining an alert, attentive state, which Colombo (2001) argued is related to the *alertness* function of attention. Changes in this function are related to the amount of time infants are in an awake alert state and reflect noncortical developmental changes (see Colombo, 2001, for a review). It is plausible that changes in infants' regulation of their state (e.g., awake and alert, drowsy, asleep) contribute to alerting as defined by Posner and Petersen (1990). Indeed, Posner and Rothbart and their colleagues have argued that behavioral regulation—and executive attention—are related developmentally to the alerting and orienting network (Posner & Rothbart, 2009; Sheese, Rothbart, Posner, White, & Fraundorf, 2008). But it is difficult to determine how visual attention versus other more general aspects of nervous system regulation determines how much of the time young infants spend fixating a stimulus.

Moreover, studies in the 1960s and 1970s on infants' visual attention focused on stimulus properties that elicit sustained attention (Fantz & Nevis, 1967). Indeed, this emphasis and body of literature led to theories such as Cohen's (1973) highly influential two-process theory of infants' visual attention. Cohen argued that how quickly young infants orient (attention-getting) to a stimulus is related to the physical properties of the stimulus (e.g., its size) whereas how long infants continued to look at a stimulus (attention-holding) is related to its complexity or how difficult it was for infants to process, form a memory, and the like. The relation between visual attention and aspects of processing or one's ongoing cognitive goals has for decades been a focus of research on visual attention across the life span (Desimone & Duncan, 1995; Folk, Remington, & Johnson, 1992; Lavie, Hirst, de Fockert, & Viding, 2004). These questions remain at the forefront of the study of visual attention. However, as we discuss later, the developmental science community now recognizes that they reflect interactions between attention and other psychological processes rather than solely visual attentional processes.

It is also important to note that the terms "attention" and "looking" historically were used interchangeably. Although the conflation of these constructs is intuitive, looking time is not the same as attention. Looking is a very gross metric of attention per se and likely reflects a conglomeration of other processes, for example, processing or learning rates, memory, and visual preference. Disentangling visual attention and looking has been difficult because of a lack of measurement tools. Historically, researchers could measure only coarse aspects of infants' looking behavior—evaluating the direction of the eyes (and head) to determine whether infants looked at a particular image, object, or person, and how long infants continued to look at an item once fixated. Developments in eye tracking (Gredebäck, Johnson, & von Hofsten, 2010) and event-related potential (ERP) methods (Reynolds, Guy, & Zhang, 2010; Richards, 2001) have opened new possibilities for examining infants' attention. In particular, such methods provide insight into infants' covert attention shifting. For example, it is now possible to determine whether infants more quickly fixate a validly cued location than an invalidly cued location (Markant & Amso, 2015; Ross-Sheehy, Schneegans, & Spencer, 2015). By measuring where and how quickly infants orient to an object or location, we can establish whether infants look more quickly at a target appearing at a cued location than at a target appearing at an uncued location, for example. If this pattern emerges, we conclude that infants must have shifted their attention to the cued location before making an eye movement; thus, such effects provide evidence of covert attentional shifts. Other work has examined the neural circuitry supporting covert attentional shifts using ERP methods (Richards, 2000, 2005).

Richards and colleagues (Richards & Casey, 1992; Richards, 1989) measured heart rate variability in young infants as a physiological index of attentional engagement during periods of looking. Specifically, Richards and Casey (1992) described heart rate defined phases of attention during periods of sustained looking at dynamic, complex video clips (e.g., moving shapes, clips from *Sesame Street*). Infants' heart rates undergo a predictable and systematic pattern of changes during looks to visual stimuli, indicating changes in the infants' level of attention engagement. Specifically, soon after initiating a fixation of a stimulus, infants' heart rates begin to decline, indicating that they are entering a state of sustained attention, where infants are found to be more resistant to distraction. After a period of sustained low heart rate, infants' heart rates increase and return to the prestimulus level, indicating sustained attention termination. These data suggest that at least by 8 weeks of age, infants' sustained fixations actually reflect several phases and that only some proportion of individual looks reflects the kinds of attentional processes discussed in the context of other procedures, at other ages, and so on. Because the stimuli used in this research are complex and often multimodal (e.g., several studies used clips from *Sesame Street*), we must be cautious about concluding that the observed patterns reflect only visual attentional processes; as with much infant work, the findings may reflect a combination of visual attentional processes in conjunction with other perceptual and cognitive processes, such as visual perceptual skill control over eye movements, learning, and memory.

A larger literature has been devoted to developmental changes in aspects of looking behavior that reflect spatial orienting processes. A primary focus has been to understand changes in voluntary control over visual attention in the first 12 postnatal months (see Ruff & Rothbart, 1996, for a review). Specifically, several researchers have concluded that attention in very young infants is *stimulus bound*, or externally controlled (Colombo, 2001); it has even been stated that their attention is *obligatory* (Stechler & Latz, 1966). These conclusions are based on the observation that in the first postnatal weeks, infants seem to be unable to disengage attention from a fixated stimulus in order to fixate another stimulus. In the gap-overlap task—in which a peripheral stimulus is presented when the infant is fixating a central stimulus— fixations of very young infants' appear to be *sticky*. In this task, infants look at a central stimulus, which then disappears and is followed by a peripheral stimulus to either the left or the right of center. (See Figure 1.4.) Reaction times to orient to the peripheral stimulus indicate infants' ability to flexibly shift orienting. In *overlap* trials, the central stimulus—the target the infant is fixating—remains visible when the peripheral stimulus is presented. Under these conditions, young infants have significant difficulty disengaging from that central stimulus and shifting their fixation to the peripheral target (Hood & Atkinson, 1993). Because, as described earlier, looking behavior is thought to reflect attention, the conclusion has been that this apparent stickiness arises from infants' inability to voluntarily shift the direction of their attention.

At about 4 months, there appears to be a shift in this "stickiness" in infants' looking behavior. Smooth pursuit rapidly develops from birth to 4 months, and at 4 months smooth pursuit dominates visual tracking (Rosander, 2007). In the overlap task just described, infants more easily shift attention by 4 months (M. H. Johnson, 1995). Recall, however, that our understanding of visual attention in infancy reflects our evaluation of

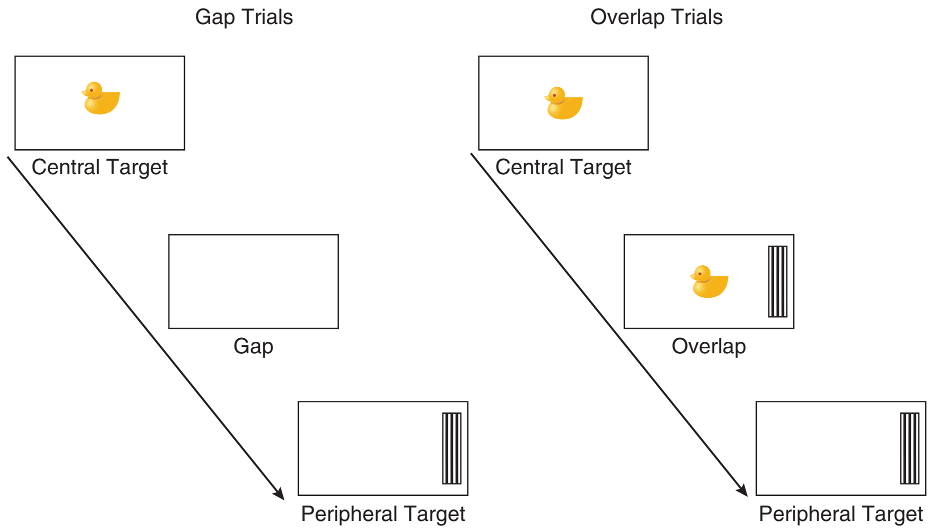


Figure 1.4 A schematic depiction of a gap-overlap task. There are two trial types: Each trial begins with a central target presented at fixation (the duck here); after some period of time the central target disappears and a peripheral target (the black and white bars here) appears. The difference between the two types of trials is whether the two targets are presented at the same time (in overlap trials) or separated by a brief blank screen (in gap trials). Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

visual behavior. Between birth and 4 months of age, there are significant changes in oculomotor control, and as a consequence, at 4 months, infants have sufficient control over eye movements such that they are reliable research participants. Although there have been discussions about the role of attention in oculomotor control (e.g., Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999) and saccadic eye movements (Canfield & Kirkham, 2001; Hoffman & Subramaniam, 1995), there is evidence that even in adults, performance on some attention tasks requiring eye movements involves multiple neural systems and does not reflect solely attentional processes. (See, e.g., Csibra, Johnson, & Tucker, 1997.) We therefore must be cautious when drawing conclusions about infants' attention from behavior that taxes oculomotor control (Nakagawa & Sukigara, 2007).

The change at 4 months in infants' ability to shift their attention in the overlap task does not mean that this aspect of visual attention is fully developed. In the second half of the first postnatal year, infants' ability to shift attention in this context varies as a function of the content of the central, fixated stimulus (Peltola, Leppänen, Palokangas, & Hietanen, 2008). This variation in the second half of the first year perhaps reflects the fact that infants' processing of the meaning or significance of the central stimulus influences their ability to detect or respond to a peripheral or distracting stimulus. In a very different context, Oakes and colleagues (2002) observed that when playing with toys, 10-month-old infants are less easily distracted by an external stimulus when they are engaged in deeper processing of those toys than when they are less engaged. At 6 months, infants show similar levels of distraction in different states of engagement, suggesting that infants' ability to control their attentional focus—and resist distraction during information processing—shows developmental change during this time.

The literature just described suggests important development in the spatial orienting of attention during the first postnatal year. More precise understanding of this development derives from work using tasks that are more closely related to the tasks developed for older populations. Specifically, a number of studies have used tasks that allow more sensitive measures of spatial orienting that are not conflated with measures of looking. These studies use a task like that illustrated in Figure 1.5. In this task, infants first are induced to fixate a central location (e.g., an interesting stimulus is presented in this location). Next, as infants fixate this centrally presented item, a peripheral cue is briefly presented to the left or right of fixation. Finally, a visual target is presented either in the validly cued location (i.e., where the cue appeared when the infant was fixating the central stimulus) or in an uncued or invalid location (i.e., on the side opposite to where the cue appeared).

Studies using this procedure have documented that visual attention orienting is facilitated to the cued location relative to the uncued location if the interval between cue and target is short. That is, the subject will detect, perceive, and process the target faster or better if it is presented in a validly cued location than if it is presented in a location that is not cued (Carrasco, 2014). Adapting this procedure for use with infants, Johnson, Posner, and Rothbart (1994) observed adult-like responses in such a task by 4-month-old infants. Infants, like adults, responded more quickly to a target that appeared in a cued location. Ross-Sheehy and colleagues (2015) recently introduced an adaptation of this method in which infants are exposed to a variety of cue conditions (e.g., validly cued targets, invalidly cued targets, and neutrally cued targets). Ross-Sheehy et al. observed that older infants showed more effective use of the cues than did younger infants, experiencing less competition between irrelevant cues.

However, spatial cuing does not always result in facilitated or faster response to the cued location. Critically, when the delay between the cue and the target is long (e.g., > 200 ms), people are actually worse at responding to a target that appears in the cued location relative to a target that appears in the uncued location. This effect, termed "inhibition of return" (IOR), presumably reflects the system inhibiting returning attention to a previously attended location. That is,

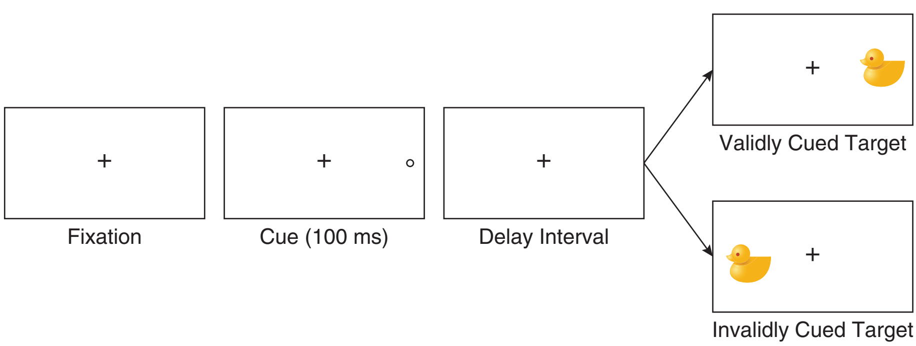


Figure 1.5 An illustration of spatial cueing attention task. When the infant is fixating the central target (the fixation cross), a cue is briefly presented in the periphery. Following a brief delay, in validly cued trials (the top frame), the target is presented in the same location as the cue. In invalidly cued trials (the bottom frame), the target is presented in the opposite location from the cue. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

IOR has been described as an adaptation of attentional mechanisms such that once a location is attended and no target occurs, the system inhibits that location in order to encourage orienting to new locations (Klein, 2000). As a result of inhibiting the cued location during the delay, any target that is presented in the cued location is also inhibited, resulting in slower eye movements to that item. There is evidence of IOR in newborns (Simion, Valenza, Umiltà, & Barba, 1995; Valenza, Simion, & Umiltà, 1994) when they are allowed to make overt shifts of attention to the cue. However, when the cue is too rapid and only a covert attention shift can be made, IOR appears to emerge at 5 to 6 months of age (Richards, 2000) and is stable by 9 months (Markant & Amso, 2013, 2015). Richards (2000) paired this task with presaccadic ERPs to show more cortical involvement of parietal and frontal sites with behavioral developmental change from infants 3 to 7 months old. This task, therefore, is a critically important addition to the available tools to assess visual attention. It offers insight into inhibitory processing, an important component of distractor suppression during target selection.

Another task that also provides insight into these inhibitory processes is the *negative priming* task. In this task, a target and a distractor initially are presented together, presumably eliciting attention to the target and inhibition to the location of the distractor. (Maintaining attention to the target presumably requires inhibiting the distractor.) Then, during a second or probe display, the target is presented alone, either in a novel (previously empty) location or in location previously occupied by the distractor. Because the location previously occupied by the distractor was ignored or inhibited, the reasoning is that infants will have more difficulty orienting to a target presented in that location. Indeed, consistent with the data from studies using IOR tasks, infants' responses to targets appearing in previously inhibited locations is slowed compared to their responses to targets appearing in previously empty locations. Thus, performance on these tasks can be used to draw conclusions about infants' ability to inhibit attention to a particular location. Moreover, the developmental changes in this task converge with those obtained when using IOR; infants show developmental change in inhibitory processing across the first postnatal year, with 3-month-olds showing no sign of inhibition but rather facilitation and with inhibitory processing being robust by 9 months (Amso & Johnson, 2005, 2008; Nakagawa & Sukigara, 2007).

Other work has attempted to evaluate infants' visual selective attention orienting more broadly by assessing their performance on visual search tasks. For example, visual search requires shifting attention to a target and inhibiting attending to distractors. A hallmark of effortful visual search is that target identification takes longer with increasing numbers of distractors—because the viewer must attend to individual items or regions of space that contain items, the more items there are, the longer (on average) it will take to find the target. (See discussion in the previous section, "Influential Models and Common Tasks.")

Variations of visual search tasks have been used to study visual attention processes in infants. Very early in infancy, we can ask what stimulus features automatically capture attention by examining visual pop-out. For example, Dannemiller (2005) observed 2-month-old and 4.5-month-old infants' orienting to a singleton oscillating target in a field of static bars. The moving target should capture infants' attention, and their ability to fixate the target and inhibit looking at the nonmoving distractors provides insight into the nature of their visual attention processing. Dannemiller found the pop-out effect in 4.5- but not yet in 2-month-old infants.

Using eye tracking, Amso and Johnson (2006) observed that 3-month-old infants effectively selected both a moving target in a field of nonmoving targets and an oriented bar in a field of vertical bars more often than would be expected by chance. Performance on the moving target search was significantly better than on the more difficult orientation-based search. Frank, Amso, and Johnson (2014) showed developmental improvement in both search tasks from 3 to 10 months of age.

Adler and Orprecio (2006) provided additional evidence that at least some aspects of visual search in infancy are similar to those in adults. They presented 3-month-old infants with two types of visual search arrays: one that should elicit a preattentive target detection for adults (detecting a + in an array of Ls, or target-present arrays) and another that should be elicit more effortful attention (an array of all Ls, or target-absent arrays). Indeed, Adler and Orprecio observed that both 3-month-old infants and adults had similar latencies to find the + in the target-present trials regardless of the number of distractors, but their performance varied considerably by the number of items in the target-absent trials. Similar results were reported by Adler and Gallego (2014). Thus, although we must be cautious about concluding that similar patterns of behavior in infants and adults necessarily reflect the same underlying processing (particularly as adults are given instructions in this task and infants are not), these findings show some similarities in how infants and adults search for targets in cluttered visual arrays.

Work using computational modeling provides insight into the developmental mechanisms behind this development, in particular the neural development that may support developmental changes in orienting during visual search early in infancy. Specifically, work using computational modeling has identified increases in the size of horizontal connections in primary visual cortex and the duration of recurrent posterior parietal activity as critical to effective visual attention orienting performance in infant visual search data (Schlesinger, Amso, & Johnson, 2007, 2012).

In a different type of visual search experiment, Kwon et al. (2016) presented 4- to 8-month-old infants with an array of 6 different photographs of familiar items (shoe, flower, vehicle). One item in each array was a human face. Whereas 4-month-old infants were drawn to the most physically salient item in the array (as defined by brightness and orientation), 6- and 8-month-old infants looked at the human faces. Studies like these uncover spontaneous behavior by infants when presented with visual search arrays and begin to reveal how infants' looking behavior (and visual attention) is controlled by external stimulus factors (such as movement or physical salience) versus other, nonphysical features (such as familiarity or meaning). Consistent with other work examining visual attention in infancy, the results of Kwon et al. showed that by 6 months, infants could use top-down content, such as familiarity or meaning inherent in a human face, to endogenously guide visual attention orienting in the presence of distraction. Data like these are consistent with the general conclusion that processes engaged in voluntary control of attention increase during the first postnatal year.

Recall that Colombo (2001) described two orienting functions of attention, one based on location and the other based on object features. As just described, most of the work on visual attention in infancy has focused on the spatial orienting function of attention. But there is a small emerging literature on object-based attention in infancy. The term “object-based visual attention” refers to attention to one of many features or objects at a particular location at the expense of others. Using cuing methods, adults have been shown to have object-based attention. For example, Egly, Driver, and Rafal (1994) presented a cue on a part of an object; this cue helped adults attend to the object, facilitating their detection of a target that subsequently is presented on that object compared to an equally distant target presented on a different object.

Bulf et al. (2013) used a variation of this task to examine object-based attention in infants. (See Figure 1.6.) In this variation, infants first saw two identical bars for a brief period of time. Then, a cue appeared on one of the two bars. After a delay (200 ms interstimulus interval [ISI] in Figure 1.6), infants then saw a target presented in the cued location or in one of two uncued locations—both equally distant from the cue. However, one kind of the uncued items (the “Invalid same-object” array in the figure) was presented on the cued object, whereas the other kind of uncued item (the “Invalid different-objects” array in the figure) was presented on the other object. Eight-month-old infants also showed object-based attention cuing benefit; they were faster to detect targets in the same-object displays relative to targets in the between-objects displays. (See also Valenza, Franchin, & Bulf, 2014.) In general, researchers agree that object-based attention effects depend heavily on the strength of object representation and recognition as well as object characteristics, such as goodness (Chen, 2012). Although object-based attention is not yet well studied in developmental science, the study of object perception and recognition enjoys a long history of developmental research beginning with Piaget.

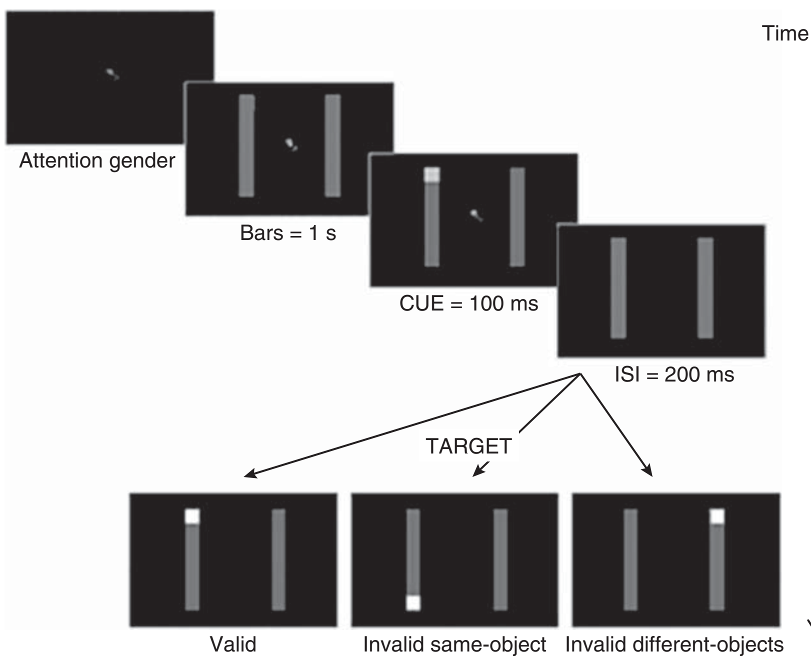


Figure 1.6 Illustrates the procedure used by Bulf & Valenza (2013) to examine object-based visual attention in 8-month-old infants.
Source: Bulf & Valenza (2013), published by APA. Reprinted with permission.

Thus, future research may build on this foundational work on infants' object-based attention and work on object perception and recognition to provide deeper insight into the development of attention more broadly.

Attention in Early Childhood

The transition from infancy to early childhood comes with continued development of visual attention processes. Notably, the relevant changes are not solely in visual attention processes. These processes in childhood operate in a different body than they had in infancy. Young children are mobile, willful, and have strong emerging language skills. Thus, visual attention processes become integrated into a larger set space of competing exploratory skills. It follows that while both alerting and orienting show some measurable developmental change into childhood, it is the executive processes that become a critical component of managing or regulating the now-dynamic opportunities facing the growing child.

Although not explicitly focused on understanding visual attention per se, early studies of toddlers' and preschool children's sustained attention during television watching provide some insight into attentional abilities, at least in the context of watching a complex, dynamic, multimodal stimulus. The findings suggest developmental changes in the *alerting* network during this period. For example, children's attention to television programming increased between age 1 and 4 years (Anderson & Levin, 1976), and children's sustained attention during television viewing was related to their comprehension of the content (Lorch, Anderson, & Levin, 1979). Such findings provide a foundation for understanding how children's sustained attention develops during early childhood and suggests that, as with infants (e.g., Cohen, 1991), the duration of periods of sustained attention is related to children's processing of the stimulus content. Moreover, 5-year-old children are less distractible—and presumably more engaged—when the content being viewed is comprehensible than when it is not (Lorch & Castle, 1997). During the preschool years, there continue to be developmental changes in children's ability to maintain an alert and engaged attentional state, and this ability is enhanced by their ability to understand the content of the stimulus being visually attended.

In addition, the study of children's general attention processes while viewing television—and to some extent during toy play—led to conclusions about the development of attentional inertia, or the process by which attention becomes more engaged over time (Richards & Anderson, 2004). The notion is that sustained attention builds and engagement with the stimulus deepens over the period of sustained attention. This conclusion is supported by the fact that children become less easily distracted as a period of sustained attention continues (Anderson, Choi, & Lorch, 1987; Oakes, Ross-Sheehy, & Kannass, 2004) and by physiological changes, including reductions in heart rate, that occur over prolonged periods of sustained attention (Richards & Cronise, 2000; Richards & Gibson, 1997). This characteristic of increasing engagement over periods of sustained attention is not specific to the preschool years; there is evidence of this process in infancy (Oakes et al., 2004) through the preschool period (Richards & Cronise, 2000). Of course, developmental changes in attention occur during this time period. Given the same stimulus, periods of sustained attention increase over age, and comprehension appears to have an increasing influence on children's sustained attention during the preschool period (Richards & Anderson, 2004).

Other work examined developmental changes in sustained attention in other contexts. For example, in a longitudinal study, Ruff et al. (1998) showed increases in children's duration of looking and focused attention between 2.5 and 4.5 years of age during free play and watching a puppet show, suggesting changes in children's ability to sustain an engaged attentional state. Moreover, the context—particularly the number of toys present—may influence whether sustained attention increases or decreased from infancy through the preschool years (Ruff & Capozzoli, 2003). Such effects underscore the close connection between attention and other cognitive processes and how attention is differentially engaged depending on the cognitive load imposed by the task. During the preschool period, there appear to be changes in the level of engagement during attention, with older children being more resistant to distraction than younger children are during periods of sustained attention during toy play (Ruff & Capozzoli, 2003). Taken together, this research has revealed changes during early childhood in the duration and the level of engagement during periods of sustained attention. Because sustained attention is related to information processing—and the comprehensibility and complexity of the stimulus content—developmental changes must be evaluated taking into consideration the nature of the stimuli, task, context, and other factors.

The work during early childhood also reveals changes in orienting. For example, Gerhardstein and Rovee-Collier (2002) used a visual search task (their stimuli are illustrated in Figure 1.2) to examine orienting in children between 1 and 3 years of age. In this task, children were taught to touch the target. Recall that in Figure 1.2a, the target is different from the distractors only by a single feature, and therefore the feature task should be easy and not require attention. Recall that the target in Figure 1.2b is defined by a conjunction of features—it is the instance that is defined by a specific color/shape combination—and search for this target should require attention and should be effortful. Indeed, Gerhardstein and Rovee-Collier found the number of items in the arrays in the feature task had no effect; the only significant effect was that younger children were slower to find the target. Thus, detecting the target did not appear to require effortful attentional orienting. In contrast, children's performance in the conjunction task varied with the number of distractors—children had more difficulty identifying the target when there were more distractors. In both tasks, younger children were generally less efficient and less accurate than older children, but the effect of attention seemed to be the same across this age range, suggesting that the only developmental effects observed here were those that reflect developmental change in young children's *general* attentional abilities, or something related to making a response. Scerif and colleagues (2004) observed similar results in a touch-screen visual search task with children in this same age range. However, because Scerif et al. also included some of the displays without targets, they could examine not only search times but also other variables, such as search paths and perseverative errors to nontargets. The inclusion of such variables may have yielded more sensitivity to developmental differences in this age range. Other work using more traditional visual search tasks (pressing a key when a target is found within an array) revealed developmental differences in somewhat older children (6–10 versus adults) in conjunction searches (Trick & Enns, 1998). Future work comparing different types of visual search tasks may reveal the source of such discrepancies.

Finally, the increased awareness that developmental changes in attention during early childhood reflect, at least in part, changes in executive attention or cognitive control has led to the development of new tasks to tap those developing systems. As described earlier, variations of the flanker task have been developed for use with children as young as 4 (McDermott et al., 2007). This task, which is depicted in Figure 1.3, simplifies the traditional flanker task by reducing the perceptual demands of the stimuli and increases the child's ability to apply existing knowledge to their processing of the stimuli. The Track-It task developed by Fisher et al. (2013) is also argued to examine executive attention.

By manipulating features of the distractors (e.g., whether they are all the same or vary), Fisher et al. (2013) argued that this task allows assessment of endogenous and exogenous factors on children's sustained selective attention.

In summary, during the toddler and preschool years, there continue to be significant changes in attentional processes, with evidence that children are becoming increasingly more efficient in their visual attention orienting and more capable of sustained attention.

Attention in Later Childhood and Adolescence

The transition into later childhood brings modest developmental change in visual attention alerting and orienting but more robust change in executive attention. Indeed, much of the work in later childhood and early adolescence has focused on cognitive control, which is closely related—and may overlap with—executive attention.

Work with older children and adolescence suggests that there is little change in orienting attention in late childhood. Enns and Brodeur (1989) showed that 5- to 9-year-old children are more influenced by an orienting cue than are adults—both in terms of the benefit of a valid cue on their attention performance and the interference from an invalid cue. (See also Konrad et al., 2005.) However, several studies have shown that by 8 to 10 years, children's orienting is adult-like. Rueda et al. (2004) showed that in the ANT by age 10, children receive the same benefit as adults from an alerting cue. Other work has shown that visual attention orienting is adult-like by 8 to 10 years (Goldberg, Maurer, & Lewis, 2001; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; Waszak, Li, & Hommel, 2010). Using a spatial cuing task, Markant and Amso (2014) did not observe developmental change in visual attention orienting, with either facilitation- or IOR-inducing timing, in children 7 to 17 years of age, which suggests stable visual attention orienting in this age range. Thus, any changes in these attention networks beyond early childhood are subtle and much less dramatic than the development that occurs in infancy and early childhood.

In contrast to alerting and orienting, the development of executive attention processes is more protracted, with changes into adolescence Executive attention processes are involved when contexts or tasks require inhibition of conflicting or interfering sources of information in the visual environment. Resolving such conflict requires some overarching rule to guide visual attention. For example, executive attention is engaged when a target is flanked by distractors that present a conflict (see Figure 1.3)—such as when the direction of flanking arrows is different from the direction of a central arrow target. Research using tasks that require attention in the context of such conflict has revealed that executive attention is not yet adult-like in childhood (Goldberg et al., 2001) and continues to develop into adolescence (Konrad et al., 2005; Waszak et al., 2010).

Additional insight into the development of executive attention comes from work using the anti-saccade task (Guitton, Buchtel, & Douglas, 1985). In this task, children are taught a rule: When a cue appears on a screen, inhibit attending to it and instead orient to the opposite side of the screen. Evidence of some competence on anti-saccade tasks is available in infants (Johnson, 1995) as well as in toddlers and young children (Scerif et al., 2005). Despite these developmental changes early in childhood, as is true for other aspects of executive attention, anti-saccade development has a prolonged developmental time course, becoming adult-like by roughly 14 years of age (Luna, Garver, Urban, Lazar, & Sweeney, 2004).

Moreover, neuroimaging data have exposed the neural networks underlying these visual attention processes; these findings confirm and provide additional insight into the behavioral changes described. For example, Konrad et al. (2005) showed that 8- to 12-year-old children had less activation than did adults in frontal-midbrain regions during alerting, less activation in the temporoparietal junction during orienting, and less activation in the dorsolateral prefrontal cortex during executive attention tasks. Using anti-saccade, Luna and colleagues (2004) have shown that developmental change in top-down executive control of visual attention involves frontoparietal engagement and emerging long-range connections between these regions and develops into adolescence (Crone, 2009; Hwang, Velanova, & Luna, 2010).

EMERGING TRENDS AND FUTURE DIRECTIONS IN THE STUDY OF THE DEVELOPMENT OF VISUAL ATTENTION

As illustrated by the preceding discussion, much work in the study of the development of visual attention has focused on demonstrating the state of the system at different developmental points. This has been a fruitful approach and has yielded significant understanding of both the limitations and the capabilities of visual attention across development.

With this work as a foundation, two trends have emerged in the literature that have and will continue to shape our understanding of the development of visual attention. The first emerging trend derives from the fact that the process-oriented focus in the study of attention has highlighted the connections between attentional processes and other processes, in particular learning and memory. Second, there has been an explosion of new tools available for studying attention. Many of these tools are further refinements of older tools or involve the application of tools used with adults or in neuropsychological contexts. However, the availability of new imaging techniques—as well as methods for analyzing the data from those techniques—has yielded significant insight into how developing neural structures influence attentional processes. We discuss each of these trends in the following paragraphs.

Attention and Its Interactions with Learning and Memory

Attention as a process interacts with learning and memory processes in intimate and complex ways. Historically, researchers have asked how cognitive processes influence attention—for example, how children are more engaged and less distractible when attending to content they understand than when attending to content that is more difficult to understand (Lorch et al., 1979). However, it is important to keep in mind that one part of the definition of attention is that, because it functions to filter distraction, it increases the efficiency of other cognitive processes. An emerging trend in the literature is a deep recognition of this connection. For example, attentional processes may differ depending on the content of information in particular location. Also, and relatedly, visual attention dynamics—such as the facilitation of processing of some information or such facilitation in combination with inhibition of competing distractors—can impact how attended items are learned and remembered.

One way attention is related to learning and memory is that how perceivers distribute their attention to a stimulus can determine what they learn about that stimulus. For example, using eye tracking, researchers have asked how people distribute their attention to specific parts of a visual scene (ignoring other parts of the scene) and how that pattern of attention relates to their learning about the objects in those scenes. These relations have been demonstrated even in infancy. Johnson, Slemmer, and Amso (2004) found a relation between where infants oriented on a visually ambiguous display (a rod divided by a central box) and whether infants perceived the central rod object in the display as complete or broken. Infants who oriented to (looked at) the object parts and their movement perceived the rod and box in an adult-like manner, whereas those who oriented randomly did not. This and other work has collectively identified a role for efficient attention-guided orienting in bootstrapping both object and face perception (Amso, Fitzgerald, Davidow, Gilhooly, & Tottenham, 2010; Amso & Johnson, 2006; Emberson & Amso, 2012; Johnson et al., 2004).

Moreover, at least in infancy, previous learning can shape how viewers orient to a stimulus, presumably influencing what they learn about those stimuli. For example, 4-month-old infants who live with a pet distribute their looking differently to images of dogs and cats than do infants who do not live with pets (Hurley & Oakes, 2015; Kovack-Lesh, McMurray, & Oakes, 2014; Markant & Amso, 2015). Similarly, infants looking at faces distribute their visual attention differently when viewing relatively familiar, own-race faces than when viewing relatively unfamiliar, other-race faces (Xiao, Quinn, Pascalis, & Lee, 2014; Xiao, Xiao, Quinn, Anzures, & Lee, 2013). Thus, not only does orienting influencing learning in the moment, but the strategies that infants use to guide their attention to a stimulus reflects their past experience.

In addition, visual attention can bias what infants learn about available content. As noted, IOR emerges by the time infants are 5 to 6 months of age. Recent work has identified a novel role for IOR, during visual attention orientating, in learning and memory (Markant & Amso, 2013, 2014; Markant, Oakes, & Amso, 2015; Markant, Worden, et al., 2015). Using spatial cuing tasks, these studies showed a benefit for objects that were attended to and encoded in the IOR condition. Recall that IOR is elicited when subjects are cued to a location, but there is a relatively long delay between the offset of the cue and the onset of the target. On these trials, participants simultaneously suppress or inhibit the cued location—that is, the distractor location—and increase attention to the noncued location—that is, the target location. Studies with infants show that when the timing elicits IOR, infants more effectively learn objects presented in the noncued (target) location than objects in the cued (distractor) location, illustrating suppression of the object in the distractor condition and facilitation of attention and learning to the object in the target location.

Moreover, functional magnetic resonance imaging (fMRI) data in adults showed that this memory benefit was linked to attentional modulation of visual cortex activity: Recognition accuracy for objects encoded in the context of IOR was predicted by cortical activity associated with target location enhancement and by the extent to which competing distractor locations were inhibited during initial encoding (Markant, Worden, et al., 2015). These data suggest that, in filtering distraction, visual attention provides a less noisy representation of the attended item for learning and memory.

Markant, Oakes, and Amso (2015) provided a powerful demonstration of this effect. They observed that they could influence how infants processed items within a category of objects by biasing infants to attend to that category. A number of studies have shown that infants orient attention differently to informative parts of own-race versus other-race faces, in particular the eyes (Wheeler et al., 2011; Xiao et al., 2013). Markant, Oakes, and Amso (2015) asked a different question; they asked how biasing infants to attend to some types of stimuli (but not to other types of stimuli) could influence asymmetries in processing faces based on race. They used a spatial cuing procedure to bias Caucasian 9-month-old infants to attend to either own- or other-race faces. (All infants were exposed to the same own- and other-race faces; some infants were biased to attend to the own-race faces, and other infants were biased to attend to the other-race faces.) Infants showed stronger discrimination of and memory for faces from the race that was the focus of the attention bias, regardless of whether those faces were from their own-familiar race or a different, unfamiliar race. Thus, the extent of attention engagement, and distractor suppression, at initial stimulus encoding—not the familiarity of the race—determined the asymmetry in processing in this case. These results extend other attentional explanations of the other-race effect in both adult (Hills, Cooper, & Pake, 2013) and infant literatures (Wheeler et al., 2011; Xiao et al., 2013) on how attention (as measured by eye movements) is distributed to different facial areas influences the other-race effect.

Finally, research using very different methods and procedures has also shown that attention can contribute to what infants learn in cluttered visual scenes. Specifically, when presented with an array of multiple objects, learning about any individual object requires selecting that objects, attending to it, and inhibiting distracting objects. This may be especially difficult for young infants. Ross-Sheehy and colleagues (2011) observed that facilitating young infants' attention to an individual item in a multiple item array allowed them to encode that individual item into visual short-term memory (VSTM) and detect when it changed. Importantly, this effect was observed at a point in development when infants appear to be unable to encode or store in VSTM individual items in multiple-item arrays (Oakes, Baumgartner, Barrett, Messenger, & Luck, 2013; Ross-Sheehy, Oakes, & Luck, 2003). These relations appear to continue across development. For example, Astle, Nobre, and Scerif (2012) observed that individual differences in attentional control were related to VSTM in 7- to 10-year-old children, providing support for the idea that developmental changes in attentional control contribute to developmental changes in VSTM. In sum, attention and memory are reciprocally interactive, and a great deal is gained by examining their development as such.

New Tools to Study Visual Attention

Behavioral Tools

Many new behavioral tools have been developed to study attentional processes in infancy and childhood. The availability of eye trackers with adaptations for calibrating and tracking younger children's eye gaze have opened the door for the introduction of new tools as well as the refinement of existing tools.

Consider the visual search tasks described throughout this chapter. Examining visual search with young children was extremely difficult until new technical tools were developed. For example, Dannemiller (2000) drew conclusions about the role of external stimulus factors on young infants' attention by examining gaze shifts to displays containing a number of static shapes and one moving shape. Using classic forced-choice preferential looking (FPL) procedure (Teller, 1979), observers watched infant behavior and made a judgment (based on head movement, eye direction, facial expression, and other idiosyncratic behaviors) about the side of the moving bar. Because the observers have no information about where the moving bar is, the observer will be accurate (i.e., be able to judge correctly the side of the display containing the moving bar) only if the infant has a strong tendency to look at that bar. This procedure has been extremely successful at evaluating many aspects of young infants' visual behavior (Dannemiller, 2000; Powers, Schneck, & Teller, 1981; Wattam-Bell, 2001), but it allows only a crude measurement of where infants are looking. Thus, it is less useful for assessing complex attentional processes in visual search.

Others have attempted to understand how attention is deployed and used in visual search with habituation or familiarization tasks (Quinn & Bhatt, 1998) or conditioning tasks (Rovee-Collier, Hankins, & Bhatt, 1992). However, these tasks also do not allow evaluation of attentional processes on the same timescale as in traditional visual search (i.e., on a single brief exposure to a stimulus array), and they also do not allow precise measurement of where subjects look.

The availability and accessibility of eye tracking systems that can be used with young children and infants has allowed researchers to ask more sophisticated questions about visual search in these age groups. Specifically, researchers can now measure, with extreme precision, exactly where infants look, how many targets they orient to prior to landing on the target, their scan paths when distractors are nearby, and the latency in milliseconds to target identification. The development of eye tracking methods has given scientists the ability to unconfound visual attention processes from other variables involved in looking behavior.

As discussed throughout this chapter, the development of new tasks also has advanced visual attention research. Tasks have been developed that are explicitly linked to Posner and Petersen's separable networks and the ANT. Of course, the most influential task is the ANT itself (Rueda et al., 2004), which has been used to assess attentional processes in children. Results from this task have shown how attentional abilities are related to executive control and emerging self-regulation (Rueda, Posner, & Rothbart, 2005). Many tasks, such as the NIH Toolbox for the Assessment of Neurological and Behavioral Function (Zelazo et al., 2013) and the Early Childhood Attention Battery (ECAB), developed by Breckenridge, Atkinson, and Braddick (Atkinson & Braddick, 2012; Breckenridge, Braddick, & Atkinson, 2013), have examined these types of relations to assess different aspects of attention in early childhood—particularly those related to executive attention and cognitive control—that are predictive of atypical developmental trajectories. For example, the ECAB has revealed deficits in attentional processes of children with Down syndrome and Williams syndrome (Breckenridge, Braddick, Anker, Woodhouse, & Atkinson, 2013) and may help both understanding and early identification of such disorders (Atkinson & Braddick, 2012).

In addition to these broad tasks, other tasks have been developed to assess specific aspects of visual attention. Ross-Sheehy and colleagues (2015) developed an attentional cuing task for use with infants and young children that takes advantage of infants' and young children's interest in moving, dynamic stimuli and presents young children with several types of cuing. In each trial, an attractive central stimulus (a looming smiley face) is presented. As infants fixate that stimulus, a cue is presented—a single cue in one of two peripheral locations, a neutral cue, or a tone—then, after a brief delay, a target is presented in one of the peripheral locations. By comparing how quickly infants fixate the target in different cuing conditions, Ross-Sheehy et al. have identified different attentional profiles in infancy and have examined developmental changes in how effective infants are in controlling their attention. Similarly, Markant and Amso (2013, 2015) have adapted a spatial cuing paradigm to examine IOR in infancy. Although IOR has been studied in infants—in particular, to document whether IOR exists in infancy (Butcher, Kalverboer, & Geuze, 1999; Valenza et al., 1994; Varga, Frick, Kapa, & Dengler, 2010)—Markant and Amso's work reflects a change in focus. As described earlier, this newer work examines the attentional processes engaged in different types of cuing and the effect of those differences on learning.

Finally, it has recently been recognized that attentional processes—particularly in infancy—can be understood through training procedures. By manipulating features of tasks and the presence or absence of reward, researchers have developed contexts in which infants and young children can be trained to use their attention in particular ways. Individual and developmental differences in how easily and effectively children can learn the contingencies and/or specific behaviors can provide insight into the systems that underlie visual attention and may help identify children at risk for developmental disorders. One such task is the Freeze Frame task developed by Holmboe and colleagues (Holmboe, Pasco Fearon, Csibra, Tucker, & Johnson, 2008; Holmboe et al., 2010b) in which children are trained to inhibit responding to a peripheral distractor. When children fixate an attractive, animated centrally presented stimulus, a peripheral stimulus is presented; when children shift their gaze to that peripheral stimulus, the central stimulus freezes. This task presumably reflects infants' emerging frontal control over visual attention, and performance at 9 months predicted performance at 24 months. Moreover, performance on this task is related to risk of later developing autism spectrum disorder (ASD; Holmboe et al., 2010a).

Similarly, Wass, Porayska-Pomsta, and Johnson (2011) found that they could use tasks like this, as well as tasks that reinforced some types of shifts of attention, to train infants' attentional control. Training had an effect on other aspects of visual attention. Specifically, training children to inhibit distractors and to follow targets increased the ability of 11-month-old infants to sustain and shift attention relative to control participants who did not receive the training.

Neuroimaging Tools

As the introduction of eye tracking technology helped bring precision the study of visual attention in infancy, so now has the introduction of novel neuroimaging tools and statistical methods provided some precision to the study of the neural architecture underlying visual attention development. Electroencephalogram (EEG) methods historically have been powerful tools for the study of the temporal dynamics of neural signals relevant to visual attention orienting (Astle, Scerif, Kuo, & Nobre, 2009; Hopf et al., 2000; Richards, 2001). One significant limitation of the EEG method is that although it has good temporal resolution, it has limited spatial resolution. Some methods have been developed to localize the source of specific ERP and EEG signals (e.g., Reynolds & Richards, 2009), but source localization of such signals is extremely coarse and subject to inaccuracies (Luck, 2014). Thus, these techniques can provide only gross indications of differences—and age-related differences—in the involvement of different neural networks during attentional processes.

The introduction of near-infrared spectroscopy (NIRS) allows better spatial precision (Aslin, 2012; Aslin, Shukla, & Emberson, 2015; Ferreri, Bigand, Perrey, & Bugaiska, 2014) and may be an essential tool for better understanding the development of cortical attention networks. NIRS uses infrared light to measure cortical activity precisely beneath the locus of the measuring optodes and emittors. The variable offered is effectively a blood-oxygen-level-dependent (BOLD) signal, which is a measurement of relative oxygenated to deoxygenated hemoglobin in response to a stimulus or event. In this way, and for the first time, the scientific community can document functional brain activations while infants are awake and performing tasks. NIRS can also be combined with tools like eye tracking to provide even more precision. Using NIRS in concert with eye tracking, for example, a recent study showed that infants engaged the dorsomedial prefrontal cortex more during a social interaction, peek-a-boo, when their partner looked directly at them rather than when the partner averted the gaze (Urakawa, Takamoto, Ishikawa, Ono, & Nishijo, 2015) Clearly, therefore, the use of NIRS is an important emerging trend in the study of visual attention, and even deeper understanding will be gained as new tasks are developed for use with NIRS. This technique also has significant limitations, however. Because the technique involves measuring how light moves through the brain, it is limited to measuring only the outermost few millimeters of the cortex.

For this reason, NIRS is unlikely to replace fMRI in child and adult participants who can perform tasks comfortably in a scanner. FMRI and diffusion tensor imaging (DTI; used to measure white matter microstructure) have long been used to study the neural underpinnings of cognitive processes in children (Amso & Casey, 2006; Qiu, Mori, & Miller, 2015). Several emerging advances mean that these techniques will be even more useful for understanding the development of visual attention networks.

Anatomical data and functional resting state data are gathered while infants and toddlers are naturally asleep. These data often can be coupled with separate behavioral data collections on tasks such as those described earlier. With respect to visual attention, for example, Elison et al. (2013) used this strategy to show that visual attention orienting as well as white matter microstructure at 7 months of age predicted later emergence of autism in an at-risk cohort. Using a similar approach, Stjerna et al. (2015) examined the relationship between visual fixations and gaze behavior and white matter microstructure at birth. Not only were better visual fixations at birth related to indices of better white matter microstructure (fractional anisotropy), visual fixation behavior related to visuocognitive task performance at 2 and 5 years of age. Thus, by combining behavioral and MRI techniques, we gain understanding into how those systems and measures are related across development.

In concert with resting-state data collection in infants and young children, advances in data analysis and modeling allow insight into the development of the neural structures that support visual attention. For example, one novel approach to fMRI research, connectomics (Di Martino et al., 2014; Sporns, 2013), has added important insight into developmental processes in particular. One view of brain development, namely interactive specialization (Johnson, 2000), holds that neural development is not the growth of any particular region but rather changes in connectivity among regions and pathways. With respect to the networks that support visual attention, connectomics data have shown that dorsal attention and frontoparietal network connectivity shows measurable strengthening even as early as 6 to 12 months (Pruett et al., 2015). Similarly, developmental improvements in executive attention are shown to be supported by strengthening of long-range connectivity from parietal to frontal regions and decreases in short-range connectivity within parietal and frontal regions (Hwang et al., 2010). These data have led many to argue that brain development is consistent with interactive specialization and that brain development involves a shift from local to long-range connections supporting increasingly mature cognition.

Semi-Naturalistic Measurement in the Study of Visual Attention

Most of the findings described here reflect results from tightly controlled, well-designed experimental procedures. The findings from such studies are invaluable for understanding cognitive processes such as those of visual attention. As evident from the preceding pages, we have gained significant understanding of visual attention across the life span using this approach. However, this approach is limited because it can provide little understanding into the ways in which visual attention operates in complex contexts, such as those encountered everyday. That is, children do not simply use visual attention to find a black circle in an array of gray squares—they use it to process the information being written on the chalkboard in a busy classroom where other children are talking, wiggling, and chewing gum, and the chalkboard is surrounded by posters, student work, and important announcements. How can we understand how the kinds of attentional processes described in this chapter translate to children's behavior in such environments?

One solution is to use semi-naturalistic measurement of visual attention, and increasingly in developmental science, methods and procedures are being developed to conduct semi-naturalistic assessments of visual attention. Technological advances allow us to take the large body of findings from tightly controlled, but relatively sparse, experimental procedures and further examine the processes using semi-naturalistic data collection techniques. One approach to semi-naturalistic data collection is to develop a laboratory space that is designed to look and function like a school classroom or room in a home. This encourages play, exploration, oculomotor engagement, locomotor action, and social interaction with others. In this way, visual attention data collection proceeds as children engage in naturalistic behaviors.

A second approach is to use high-tech solutions to systematically evaluate children's visual attention in these contexts or "in the wild," such as at home or school. For example, a number of studies have provided significant insight into attention in real-world contexts simply by recording the visual world from the infants' perspective. These studies have explored what visual information infants actually *can* attend to by simply asking what visual information is *there*. Answering this question is possible with the availability of lightweight, remote (and wireless) video recording devices that can be mounted on an infant's' forehead. Sugden, Mohamed-Ali, and Mouleson (2014), for example, placed a small camera inside an infant-size headband and asked parents of 1- and 3-month-old infants to place the headband on their infant's head whenever the infant was awake during a 2-week period. This procedure yielded hundreds of hours of recordings of what information was available to these infants, and conclusions could be drawn about how often infants could actually attend to faces in their daily lives. Such information is invaluable in understanding the real-world contexts in which infants actually use their attentional processes.

In an extension of this method, Aslin (2009) presented infants with recordings obtained from a different infant's forehead. Because the head-mounted cameras provide information only about the information infants might look at, Aslin used eye-tracking methods to record infants' eye gaze when watching the video recordings. This work shows how infants direct their gaze at scenes recorded from an infant's' viewpoint.

Such work is important for understanding how infants look at more naturalistic stimuli, but it still does not allow conclusions about how infants direct their attention during more naturalistic interactions with objects and others. That is, a key question is how infants and children deploy attention, control, inhibit, and select as they reach for objects, navigate environments, interact with others, learn the names of objects, and other activities. The development of head-mounted eye trackers has made possible the evaluation of visual attention under a range of naturalistic contexts. For example, Franchak and Adolph (2010; Franchak, Kretch, Soska, & Adolph, 2011) used head-mounted eye trackers to understand developmental changes in visual attention during developmental changes in motor abilities. Franchak and Adolph (2010) found that children and adults attended differently to obstacles as they walked around a space. Kretch et al. (2014) found that crawling and walking infants directed their gaze differently at caregivers as they approached them (e.g., crawled or walked toward them). These semi-naturalistic observations allowed researchers to understand how changes in locomotor ability—as well as age—corresponded to changes in visual attention. Similarly, Yu and Smith (2011, 2013) have used head-mounted eye trackers to examine how children's attentional processes constrain, shape, and interact with their learning of new object labels.

CONCLUSION

Visual attention is one of many attention processes that operate over sensory modalities. As is clear from the work reviewed here, a great deal of research effort has been aimed at understanding the development of visual attention beginning in infancy and at uncovering the neural mechanisms that support these processes and their development. Visual attention involves both excitatory and inhibitory mechanisms, and its development has functional significance for other cognitive and social domains in the developing child. Indeed, visual attention processes give us a window into the developing brain, are of the earliest emerging processes that are measurable in young infants, and are critical in determining what information enters the system for subsequent perception and learning. As such, visual attention processes are building block processes for perception and cognition, and their impairment has cascading effects on brain and cognitive development. The work reviewed in this chapter collectively serves an additional purpose of informing the community of scholars engaged in improving the lives of children with neurodevelopmental disorders. Visual attention processes are impaired in a variety of neurodevelopmental disorders including ASD, fragile X, and attention-deficit/hyperactivity disorder (ADHD) (see Amso & Scerif, 2015, for review).

A recent trend in the study of disorders with known impairments in visual attention is to use the described developmental trajectories of visual attention processes to predict whether an infant at familial risk for disorders will deviate from typical trajectories (Gliga, Bedford, Charman, & Johnson, 2015; Jones & Klin, 2013). For example, Elsabbagh et al. (2009) observed that infant siblings of children with ASD showed reduced attentional disengagement in comparison to siblings of children without ASD. Similarly, infants at risk for ADHD have been shown to have some differences in sensory processing as measured by ERP that later related to ADHD symptomology (externalizing behavior, attentional problems; Hutchinson, De Luca, Doyle, Roberts, & Anderson, 2013). These data provide evidence of a broader endophenotype associated with differences in visual attention modulation in infants and children at familial risk for both neurodevelopmental disorders. Thus, visual attention processes are starting to serve as biomarkers of need for early intervention.

REFERENCES

Adler, S. A. (2005). Visual search and pop-out in infancy. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *The Neurobiology of Attention* (pp. 207–212). Amsterdam: Elsevier. http://doi.org/10.1016/B978-012375731-9/50038-0

Adler, S. A., & Gallego, P. (2014). Search asymmetry and eye movements in infants and adults. *Attention, Perception & Psychophysics*, *76*(6), 1590–608. http://doi.org/10.3758/s13414-014-0667-6

Adler, S. A., & Orprecio, J. (2006). The eyes have it: visual pop-out in infants and adults. *Developmental Science*, *9*(2), 189–206.

Amso, D., & Casey, B. J. (2006). Beyond what develops when: neuroimaging may inform how cognition changes with development. *Psychological Science*, *15*(1), 24–29.

Amso, D., Fitzgerald, M., Davidow, J., Gilhooly, T., & Tottenham, N. (2010). Visual exploration strategies and the development of infants' facial emotion discrimination. *Frontiers in Psychology, 1*. Article 180. doi:10.3389/fpsyg.2010.00180

Amso, D., Haas, S., & Markant, J. (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS ONE*, *9*(1), e85701. http://doi.org/10.1371/journal.pone.0085701

Amso, D., & Johnson, S. P. (2005). Selection and inhibition in infancy: evidence from the spatial negative priming paradigm. *Cognition*, *95*(2), B27–B36. http://doi.org/10.1016/j.cognition.2004.08.006

Amso, D., & Johnson, S. P. (2006). Learning by selection: visual search and object perception in young infants. *Developmental Psychology*, *42*(6), 1236–1245.

Amso, D., & Johnson, S. P. (2008). Development of visual selection in 3- to 9-month-olds: Evidence from saccades to previously ignored locations. *Infancy: Official Journal of the International Society on Infant Studies*, *13*(6), 675–686. http://doi.org/10.1080/15250000802459060

Amso, D., & Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nature Reviews*, *16*(10), 606–619. http://doi.org/10.1038/nrn4025

Anderson, D. R., Choi, H. P., & Lorch, E. P. (1987). Attentional inertia reduces distractibility during young children's TV viewing. *Child Development*, *58*(3), 798–806.

Anderson, D. R., & Levin, S. R. (1976). Young children's attention to "Sesame Street." *Child Development*, *47*(3), 806–811. http://doi.org/10.2307/1128198

Aslin, R. N. (2009). How infants view natural scenes gathered from a head-mounted camera. *Optometry and Vision Science*, *86*(6), 561–565.

Aslin, R. N. (2012). Questioning the questions that have been asked about the infant brain using near-infrared spectroscopy. *Cognitive Neuropsychology*, *29*(1–2), 7–33. http://doi.org/10.1080/02643294.2012.654773

Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic correlates of cognition in human infants. *Annual Review of Psychology*, *66*(1), 349–379. http://doi.org/10.1146/annurev-psych-010213-115108

Astle, D. E., Nobre, A. C., & Scerif, G. (2012). Attentional control constrains visual short-term memory: insights from developmental and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(2), 277–294. http://doi.org/10.1080/17470218.2010.492622

Astle, D. E., Scerif, G., Kuo, B.-C., & Nobre, A. C. (2009). Spatial selection of features within perceived and remembered objects. *Frontiers in Human Neuroscience*, *3*, 6.

Atkinson, J., & Braddick, O. (2012). Visual attention in the first years: typical development and developmental disorders. *Developmental Medicine & Child Neurology*, *54*(7), 589–595.

Awh, E., Vogel, E. K., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208. http://doi.org/10.1016/j.neuroscience.2005.08.023

Bertenthal, B., & Von Hofsten, C. (1998). Eye, head and trunk control: The foundation for manual development. *Neuroscience and Biobehavioral Reviews*, *22*(4), 515–520. http://doi.org/10.1016/S0149-7634(97)00038-9

Breckenridge, K., Braddick, O., Anker, S., Woodhouse, M., & Atkinson, J. (2013). Attention in Williams syndrome and Down's syndrome: performance on the new early childhood attention battery. *British Journal of Developmental Psychology*, *31*(2), 257–269. http://doi.org/10.1111/bjdp.12003

Breckenridge, K., Braddick, O., & Atkinson, J. (2013). The organization of attention in typical development: a new preschool attention test battery. *British Journal of Developmental Psychology*, *31*(Pt 3), 271–88. http://doi.org/10.1111/bjdp.12004

Bulf, H., & Valenza, E. (2013). Object-based visual attention in 8-month-old infants: evidence from an eye-tracking study. *Developmental Psychology*, *49*(10), 1909–1918. http://doi.org/10.1037/a0031310

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, *33*(2), 301–311.

Butcher, P. R., Kalverboer, A. F., & Geuze, R. H. (1999). Inhibition of return in very young infants: a longitudinal study. *Infant Behavior and Development*, *22*, 303–319. http://doi.org/10.1016/S0163-6383(99)00013-2

Canfield, R. L., & Kirkham, N. Z. (2001). Infant cortical development and the prospective control of saccadic eye movements. *Infancy*, *2*, 197–211. http://doi.org/10.1207/S15327078IN0202_5

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Research*, *51*(13), 1484–1525. http://doi.org/10.1016/j.visres.2011.04.012

Carrasco, M. (2014). Spatial attention: perceptual modulation. In K. Nobre & S. Kastner (Eds.), *The Oxford Handbook of Attention* (pp. 183–230). Oxford, UK: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199675111.001.0001

Chen, Z. (2012). Object-based attention: a tutorial review. *Attention, Perception, & Psychophysics*, *74*(5), 784–802. http://doi.org/10.3758/s13414-012-0322-z

Cohen, L. B. (1973). A two-process model of infant attention. *Merrill-Palmer Quarterly*, *19*, 157–180.

Cohen, L. B. (1991). Infant attention: an information processing approach. In M. J. Weiss & P. R. Zelazo (Eds.), *Newborn Attention: Biological Constraints and Influence of Experience* (pp. 1–21). Norwood, NJ: Ablex.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367.

Crone, E. A. (2009). Executive functions in adolescence: inferences from brain and behavior. *Developmental Science*, *12*(6), 825–830. http://doi.org/10.1111/j.1467-7687.2009.00918.x

Csibra, G., Johnson, M. H., & Tucker, L. A. (1997). Attention and oculomotor control: a high-density ERP study of the gap effect. *Neuropsychologia*, *35*(6), 855–865.

Dannemiller, J. L. (2000). Competition in early exogenous orienting between 7 and 21 weeks. *Journal of Experimental Child Psychology*, *76*(4), 253–274.

Dannemiller, J. L. (2005). Motion popout in selective visual orienting at 4.5 but not at 2 months in human infants. *Infancy*, *8*, 201–216.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Di Martino, A., Fair, D. A., Kelly, C., Satterthwaite, T. D., Castellanos, F. X., Thomason, M. E., . . . Milham, M. P. (2014). Unraveling the miswired connectome: a developmental perspective. *Neuron*, *83*(6), 1335–1353. http://doi.org/10.1016/j.neuron.2014.08.050

Donnelly, N., Cave, K., Greenway, R., Hadwin, J. A, Stevenson, J., & Sonuga-Barke, E. (2007). Visual search in children and adults: top-down and bottom-up mechanisms. *Quarterly Journal of Experimental Psychology*, *60*(1), 120–136. http://doi.org/10.1080/17470210600625362

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*(2), 161–177. http://doi.org/10.1037/0096-3445.123.2.161

Elison, J. T., Paterson, S. J., Wolff, J. J., Reznick, J. S., Sasson, N. J., Gu, H., . . . Piven, J. (2013). White matter microstructure and atypical visual orienting in 7-month-olds at risk for autism. *American Journal of Psychiatry*, *170*(8), 899–908. http://doi.org/10.1176/appi.ajp.2012.12091150

Elsabbagh, M., Volein, A., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, S., . . . Johnson, M. H. (2009). Visual orienting in the early broader autism phenotype: disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, *50*(5), 637–642. doi:10.1111/j.1469-7610.2008.02051.x

Emberson, L. L., & Amso, D. (2012). Learning to sample: eye tracking and fMRI indices of changes in object perception. *Journal of Cognitive Neuroscience*, *24*(2006), 2030–2042. http://doi.org/10.1162/jocn_a_00259

Enns, J. T., & Brodeur, D. A. (1989). A developmental study of covert orienting to peripheral visual cues. *Journal of Experimental Child Psychology*, *48*(2), 171–189. http://doi.org/10.1016/0022-0965(89)90001-5

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. http://doi.org/10.3758/BF03203267

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage*, *26*(2), 471–479. http://doi.org/10.1016/j.neuroimage.2005.02.004

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. http://doi.org/10.1162/089892902317361886

Fantz, R. L., & Nevis, S. (1967). Pattern preferences and perceptual-cognitive development in early infancy. *Merrill Palmer Quarterly: Journal of Developmental Psychology*, *13*(1), 77–108.

Ferreri, L., Bigand, E., Perrey, S., & Bugaiska, A. (2014). The promise of Near-Infrared Spectroscopy (NIRS) for psychological research: a brief review. *L'Année psychologique*, *114*(3), 537–569. http://doi.org/10.4074/S0003503314003054

Fisher, A., Thiessen, E., Godwin, K., Kloos, H., & Dickerson, J. (2013). Assessing selective sustained attention in 3- to 5-year-old children: evidence from a new paradigm. *Journal of Experimental Child Psychology*, *114*(2), 275–294. http://doi.org/10.1016/j.jecp.2012.07.006

Folk, C. L., Remington, R. W., & Johnson, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030–1044.

Franchak, J. M., & Adolph, K. E. (2010). Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vision Research*, *50*(24), 2766–2774. http://doi.org/10.1016/j.visres.2010.09.024

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: a new method to describe infant looking. *Child Development*, *82*(6), 1738–50. http://doi.org/10.1111/j.1467-8624.2011.01670.x

Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, *118*, 13–26. http://doi.org/10.1016/j.jecp.2013.08.012

Gerhardstein, P., & Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *Journal of Experimental Child Psychology*, *81*(2), 194–215. http://doi.org/10.1006/jecp.2001.2649

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*(1), 1–41.

Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2015). Enhanced visual search in infancy predicts emerging autism symptoms. *Current Biology*, *25*(13), 1727–1730. http://doi.org/10.1016/j.cub.2015.05.011

Goldberg, M. C., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: the effects of endogenous cueing and of distractors. *Developmental Science*, *4*, 209–219. http://doi.org/10.1111/1467-7687.00166

Gredebäck, G., Johnson, S. P., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*(1), 1–19. http://doi.org/10.1080/87565640903325758

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research*, *58*(3), 455–472. http://doi.org/10.1007/BF00235863

Günther, T., Konrad, K., Häusler, J., Saghraoui, H., Willmes, K., & Sturm, W. (2014). Developmental differences in visual and auditory attention: A cross-sectional study. *Zeitschrift für Neuropsychologie*,, *25*(3), 143–152. http://doi.org/10.1024/1016-264X/a000126

Hills, P. J., Cooper, R. E., & Pake, J. M. (2013). Removing the own-race bias in face recognition by attentional shift using fixation crosses to diagnostic features: an eye-tracking study. *Visual Cognition*, *21*, 876–898.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795.

Holmboe, K., Elsabbagh, M., Volein, A., Tucker, L. A., Baron-Cohen, S., Bolton, P., . . . Johnson, M. H. (2010a). Frontal cortex functioning in the infant broader autism phenotype. *Infant Behavior & Development*, *33*(4), 482–491. http://doi.org/10.1016/j.infbeh.2010.05.004

Holmboe, K., Nemoda, Z., Fearon, R. M. P., Csibra, G., Sasvari-Szekely, M., & Johnson, M. H. (2010b). Polymorphisms in dopamine system genes are associated with individual differences in attention in infancy. *Developmental Psychology*, *46*(2), 404–416. http://doi.org/10.1037/a0018180

Holmboe, K., Pasco Fearon, R. M., Csibra, G., Tucker, L. A., & Johnson, M. H. (2008). Freeze-frame: a new infant inhibition task and its relation to frontal cortex tasks during infancy and early childhood. *Journal of Experimental Child Psychology*, *100*(2), 89–114. http://doi.org/10.1016/j.jecp.2007.09.004

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, *16*(4), 405–422.

Hopf, J. M., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., & Heinze, H. J. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex*, *10*(12), 1233–1241.

Hurley, K. B., & Oakes, L. M. (2015). Experience and distribution of attention: pet exposure and infants' scanning of animal images. *Journal of Cognition and Development*, *16*(1), 11–30. http://doi.org/10.1080/15248372.2013.833922

Hutchinson, E. A., De Luca, C. R., Doyle, L. W., Roberts, G., & Anderson, P. J. (2013). School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics*, *131*(4), e1053–e1061. http://doi.org/10.1542/peds.2012-2311

Hwang, K., Velanova, K., & Luna, B. (2010). Strengthening of top-down frontal cognitive control networks underlying the development of inhibitory control: a functional magnetic resonance imaging effective connectivity study. *Journal of Neuroscience*, *30*(46), 15535–15545. http://doi.org/10.1523/JNEUROSCI.2825-10.2010

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology*, *28*(5), 281–291. http://doi.org/10.1002/dev.420280504

Johnson, M. H. (2000). Functional brain development in infants: elements of an interactive specialization framework. *Child Development*, *71*(1), 75–81. http://doi.org/http://dx.doi.org/10.1111/1467-8624.00120

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1994). Facilitation of saccades toward a covertly attended location in early infancy. *Psychological Science*, *5*(2), 90–92. http://doi.org/10.1111/j.1467-9280.1994.tb00636.x

Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: eye movements and object perception performance in 3-month-olds. *Infancy*, *6*(2), 185–201. http://doi.org/10.1207/s15327078in0602_3

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–431. http://doi.org/10.1038/nature12715

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, J., . . . Fink, G. R. (2005). Development of attentional networks: an fMRI study with children and adults. *NeuroImage*, *28*(2), 429–439.

Kovack-Lesh, K. A., McMurray, B., & Oakes, L. M. (2014). Four-month-old infants' visual investigation of cats and dogs: relations with pet experience and attentional strategy. *Developmental Psychology*, *50*(2), 402–413. http://doi.org/10.1037/a0033195

Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. *Child Development*, *85*(4), 1503–1518. http://doi.org/10.1111/cdev.12206

Kwon, M.-K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2016). The development of visual search in infancy: attention to faces salience. *Developmental Psychology*, *52*(4), 537–555. doi:10.1037/dev0000080

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. http://doi.org/10.1037/0096-3445.133.3.339

Lorch, E. P., Anderson, D. R., & Levin, S. R. (1979). The relationship of visual attention to children's comprehension of television. *Child Development*, *50*(3), 722–727. http://doi.org/10.1111/j.1467-8624.1979.tb02420.x

Lorch, E. P., & Castle, V. J. (1997). Preschool children's attention to television: visual attention and probe response times. *Journal of Experimental Child Psychology*, *66*, 111–127. http://doi.org/10.1006/jecp.1997.2372

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Luck, S. J., & Vecera, S. P. (2002). Attention. In H. Pashler & S. Yantis (Eds.), *Steven's Handbook of Experimental Psychology (3rd ed.), Vol. 1: Sensation and Perception* (pp. 235–286). Hoboken, NJ: Wiley.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*(5), 1357–1372. http://doi.org/10.1111/j.1467-8624.2004.00745.x

Markant, J., & Amso, D. (2013). Selective memories: infants' encoding is enhanced in selection via suppression. *Developmental Science*, *16*(6), 926–940. http://doi.org/10.1111/desc.12084

Markant, J., & Amso, D. (2014). Leveling the playing field: attention mitigates the effects of intelligence on memory. *Cognition*, *131*(2), 195–204. http://doi.org/10.1016/j.cognition.2014.01.006

Markant, J., & Amso, D. (2015). The development of selective attention orienting is an agent of change in learning and memory efficacy. *Infancy*, *21*(2), 154–176. http://doi.org/10.1111/infa.12100

Markant, J., Oakes, L. M., & Amso, D. (2015). Visual selective attention biases contribute to the other-race effect among 9-month-old infants. *Developmental Psychobiology*, *58*(3), 355–365. http://doi.org/10.1002/dev.21375

Markant, J., Worden, M. S., & Amso, D. (2015). Not all attention orienting is created equal:

recognition memory is enhanced when attention orienting involves distractor suppression. *Neurobiology of Learning and Memory*, *120*, 28–40. http://doi.org/10.1016/j.nlm.2015.02.006

McDermott, J. M., Pérez-Edgar, K., & Fox, N. A. (2007). Variations of the flanker paradigm: assessing selective attention in young children. *Behavior Research Methods*, *39*(1), 62–70. http://doi.org/10.3758/BF03192844

Nakagawa, A., & Sukigara, M. (2007). Infant eye and head movements toward the side opposite the cue in the anti-saccade paradigm. *Behavioral and Brain Functions*, *3*, 5. http://doi.org/10.1186/1744-9081-3-5

Oakes, L. M., Baumgartner, H. A., Barrett, F. S., Messenger, I. M., & Luck, S. J. (2013). Developmental changes in visual short-term memory in infancy: Evidence from eye-tracking. *Frontiers in Psychology*, *4* (October). http://doi.org/10.3389/fpsyg.2013.00697

Oakes, L. M., Kannass, K. N. K. N., & Shaddy, D. J. J. (2002). Developmental changes in endogenous control of attention: the role of target familiarity on infants' distraction latency. *Child Development*, *73*(6), 1644–1655. http://doi.org/http://dx.doi.org/10.1111/1467-8624.00496

Oakes, L. M., Ross-Sheehy, S., & Kannass, K. N. (2004). Attentional engagement in infancy: the interactive influence of attentional inertia and attentional state. *Infancy*, *5*(2), 239–252. http://doi.org/10.1207/s15327078in0502_8

Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental Science*, *11*(1), 60–68. http://doi.org/10.1111/j.1467-7687.2007.00659.x

Petersen, S. S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. http://doi.org/10.1146/annurev-neuro-062111-150525

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25. http://doi.org/10.1080/00335558008248231

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. http://doi.org/10.1146/annurev.ne.13.030190.000325

Posner, M. I., & Rothbart, M. K. (2009). Toward a physical basis of attention and self-regulation. *Physics of Life Reviews*, *6*(2), 103–120. http://doi.org/10.1016/j.plrev.2009.02.001

Powers, M. K., Schneck, M., & Teller, D. Y. (1981). Spectral sensitivity of human infants at absolute visual threshold. *Vision Research. 21*(7), 1005–1016.

Pruett, J. R., Kandala, S., Hoertel, S., Snyder, A. Z., Elison, J. T., Nishino, T., . . . Piven, J. (2015). Accurate age classification of 6- and 12-month-old infants based on resting-state functional connectivity magnetic resonance imaging data. *Developmental Cognitive Neuroscience*, *12*, 123–133. http://doi.org/10.1016/j.dcn.2015.01.003

Qiu, A., Mori, S., & Miller, M. I. (2015). Diffusion tensor imaging for understanding brain development in early life. *Annual Review of Psychology*, *66*, 853–876. http://doi.org/10.1146/annurev-psych-010814-015340

Quinn, P. C., & Bhatt, R. (1998). Visual pop-out in young infants: convergent evidence and an extension. *Infant Behavior and Development*, *21*(2), 273–288.

Reynolds, G. D., Guy, M. W., & Zhang, D. (2010). Neural correlates of individual differences in infant visual attention and recognition memory. *Infancy*, *16*(4), 368–391.

Reynolds, G. D., & Richards, J. E. (2009). Cortical source localization of infant cognition. *Developmental Neuropsychology*, *34*(3), 312–329.

Richards, J. E. (1989). Development and stability in visual sustained attention in 14-, 20-, and 26-week-old infants. *Psychophysiology*, *26*(4), 422–430.

Richards, J. E. (2000). Localizing the development of covert attention in infants with scalp event-related potentials. *Developmental Psychology*, *36*(1), 91–108. http://doi.org/10.1037/0012-1649.36.1.91

Richards, J. E. (2001). Cortical indexes of saccade planning following covert orienting in

20-week-old infants. *Infancy*, *2*(2), 135–157. doi:10.1207/S15327078IN0202_2

Richards, J. E. (2005). Localizing cortical sources of event-related potentials in infants' covert orienting. *Developmental Science*, *8*(3), 255–278.

Richards, J. E., & Anderson, D. R. (2004). Attentional inertia in children's extended looking at television. *Advances in Child Development and Behavior*, *32*(C), 163–212. http://doi.org/10.1016/S0065-2407(04)80007-7

Richards, J. E., & Casey, B. J. (1992). Development of sustained visual attention in the human infant. In B. A. Campbell, H. Hayne, & R. Richardson (Eds.), *Attention and information processing in infants and adults: Perspectives from human and animal research* (pp. 30–60). Hillsdale, NJ: Erlbaum.

Richards, J. E., & Cronise, K. (2000). Extended visual fixation in the early preschool years: look duration, heart rate changes, and attentional inertia. *Child Development*, *71*(3), 602–620. http://doi.org/10.1111/1467-8624.00170

Richards, J. E., & Gibson, T. L. (1997). Extended visual fixation in young infants: look distributions, heart rate changes, and attention. *Child Development*, *68*(6), 1041–1056. http://doi.org/10.2307/1132290

Rosander, K. (2007). Visual tracking and its relationship to cortical development. *Progress in Brain Research*, *164*, 105–22. http://doi.org/10.1016/S0079-6123(07)64006-0

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807–1822. http://doi.org/10.1046/j.1467-8624.2003.00639.x

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2011). Exogenous attention influences visual short-term memory in infants. *Developmental Science*, *14*(3), 490–501. http://doi.org/10.1111/j.1467-7687.2010.00992.x

Ross-Sheehy, S., Schneegans, S., & Spencer, J. P. (2015). The infant orienting with attention task: assessing the neural basis of spatial attention in infancy. *Infancy*, *20*(5), 467–506. http://doi.org/10.1111/infa.12087

Rovee-Collier, C., Hankins, E., & Bhatt, R. (1992). Textons, visual pop-out effects, and object recognition in infancy. *Journal of Experimental Psychology: General*, *121*(4), 435–445.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. http://doi.org/10.1016/j.neuropsychologia.2003.12.012

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: contributions to the emergence of self-regulation. *Developmental Neuropsychology*, *28*(2), 573–594.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(41), 14931–14936.

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, *39*(5), 877–890. http://doi.org/10.1037/0012-1649.39.5.877

Ruff, H. A., Capozzoli, M., & Weissberg, R. (1998). Age, individuality, and context as factors in sustained visual attention during the preschool years. *Developmental Psychology*, *34*(3), 454–64. http://doi.org/10.1037/0012-1649.34.3.454

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in Early Development*. New York, NY: Oxford University Press.

Scerif, G., Cornish, K., Wilding, J., Driver, J., & Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Developmental Science*, *7*(1), 116–130. http://doi.org/10.1111/j.1467-7687.2004.00327.x

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience*, *17*(4), 591–604.

Schlesinger, M., Amso, D., & Johnson, S. P. (2007). The neural basis for visual selective attention in young infants: a computational account. *Adaptive Behavior*, *15*(2), 135–148. doi: https://doi.org/10.1177/1059712307078661

Schlesinger, M., Amso, D., & Johnson, S. P. (2012). Simulating the role of visual selective attention during the development of perceptual completion. *Developmental Science*, *15*(6), 739–752. http://doi.org/10.1111/j.1467-7687.2012.01177.x

Sheese, B. E., Rothbart, M. K., Posner, M. I., White, L. K., & Fraundorf, S. H. (2008). Executive attention and self-regulation in infancy. *Infant Behavior & Development*, *31*(3), 501–510. http://doi.org/10.1016/j.infbeh.2008.02.001

Simion, F., Valenza, E., Umiltà, C., & Barba, B. D. (1995). Inhibition of return in newborns is temporo-nasal asymmetrical. *Infant Behavior and Development*, *18*(2), 189–194.

Smith, N. A., & Trainor, L. J. (2011). Auditory stream segregation improves infants' selective attention to target tones amid distracters. *Infancy*, *16*(6), 655–668. http://doi.org/10.1111/j.1532-7078.2011.00067.x

Sporns, O. (2013). The human connectome: origins and challenges. *NeuroImage*, *80*, 53–61. http://doi.org/10.1016/j.neuroimage.2013.03.023

Stechler, G., & Latz, E. (1966). Some observations on attention and arousal in the human infant. *Journal of the American Academy of Child Psychiatry*, *5*(3), 517–525. http://doi.org/10.1016/S0002-7138(09)62098-7

Stjerna, S., Sairanen, V., Grohn, R., Andersson, S., Metsaranta, M., Lano, A., & Vanhatalo, S. (2015). Visual fixation in human newborns correlates with extensive white matter networks and predicts long-term neurocognitive development. *Journal of Neuroscience*, *35*(12), 4824–4829. http://doi.org/10.1523/JNEUROSCI.5162-14.2015

Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, *56*, 249–261. http://doi.org/10.1002/dev.21183

Teller, D. Y. (1979). The forced-choice preferential looking procedure: a psychophysical technique for use with human infants. *Infant Behavior and Development*, *2*(1962), 135–153. http://doi.org/10.1016/S0163-6383(79)80016-8

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1595–1608. http://doi.org/10.1037/0096-1523.25.6.1595

Treisman, A. M. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1295–1306.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trick, L. M., & Enns, J. T. (1998). Lifespan changes in attention: the visual search task. *Cognitive Development*, *13*(3), 369–386.

Ungerleider, L. G., & Pessosa, L. (2008). What and where pathways. *Scholarpedia*, *2*, 5342.

Urakawa, S., Takamoto, K., Ishikawa, A., Ono, T., & Nishijo, H. (2015). Selective medial prefrontal cortex responses during live mutual gaze interactions in human infants: an fNIRS study. *Brain Topography*, *28*(5), 691–701. http://doi.org/10.1007/s10548-014-0414-2

Valenza, E., Franchin, L., & Bulf, H. (2014). How a face may affect object-based attention: evidence from adults and 8-month-old infants. *Frontiers in Integrative Neuroscience*, *8*(March), 1–10. http://doi.org/10.3389/fnint.2014.00027

Valenza, E., Simion, F., & Umiltà, C. (1994). Inhibition of return in newborn infants. *Infant Behavior and Development*, *17*(3), 293–302. https://doi.org/10.1016/0163-6383(94)90009-4

Varga, K., Frick, J. E., Kapa, L. L., & Dengler, M. J. (2010). Developmental changes in inhibition of return from 3 to 6 months of age. *Infant Behavior and Development*, *33*(2), 245–249. doi:10.1016/j.infbeh.2009.12.011

von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences*, *8*(6), 266–72. http://doi.org/10.1016/j.tics.2004.04.002

Wass, S., Porayska-Pomsta, K., & Johnson, M. H. (2011). Training attentional control in infancy. *Current Biology*, *21*(18), 1543–1547. http://doi.org/10.1016/j.cub.2011.08.004

Waszak, F., Li, S.-C., & Hommel, B. (2010). The development of attentional networks: cross-sectional findings from a life span sample. *Developmental Psychology*, *46*(2), 337–349. http://doi.org/10.1037/a0018541

Wattam-Bell, J. (2001). The effect of contrast on vertical motion processing asymmetries in 11-week-old infants. *Perception*, *30*(2), 159–166. http://doi.org/10.1068/p3098

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own- and other-race faces differently. *PLoS ONE*, *6*(4), e18621. http://doi.org/10.1371/journal.pone.0018621

Xiao, W. S., Quinn, P. C., Pascalis, O., & Lee, K. (2014). Own- and other-race face scanning in infants: implications for perceptual narrowing. *Developmental Psychobiology*, *56*(2), 262–273.

Xiao, W. S., Xiao, N. G., Quinn, P. C., Anzures, G., & Lee, K. (2013). Development of face scanning for own- and other-race faces in infancy. *International Journal of Behavioral Development*, *37*(2), 100–105. http://doi.org/10.1177/0165025412467584

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*(2), 165–180. http://doi.org/10.1111/j.1467-7687.2010.00958.x

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE*, *8*(11), e79659. http://doi.org/10.1371/journal.pone.0079659

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox cognition battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. http://doi.org/10.1111/mono.12032

Zhang, Y., Meyers, E. M., Bichot, N. P., Serre, T., Poggio, T. A., & Desimone, R. (2011). Object decoding with attention in inferior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(21), 8850–8855. http://doi.org/10.1073/pnas.1100999108

CHAPTER 2

Category Learning and Conceptual Development

VLADIMIR M. SLOUTSKY

INTRODUCTION

Many abilities reflect the remarkable intelligence of humans: People make inferences, develop and use scientific theories, make laws, preserve knowledge and pass it onto new generations, write fiction, reason about past and future, and form counterfactual arguments. These abilities require sophisticated conceptual knowledge, much of which has to be acquired. Therefore, one of the most interesting and exciting challenges in the study of human cognition is to understand how people acquire this knowledge in the course of development and learning. In this chapter, I address this challenge and review research on conceptual development that contributes to our understanding of these issues.

What Are Concepts?

In the simplest possible way, the term "concepts" can be defined as lexicalized equivalence classes. What is an equivalence class? In his chapter focusing on concepts (Chapter XII), William James (1890/1983) wrote: "Our principle only lays it down that the mind makes continual use of the notion of sameness, and if deprived of it, would have a different structure from what it has" In other words, the mind can treat different things as if they were equivalent in some way. If the mind is capable of detecting sameness in a diverse set of objects, then a concept is an output of this process. Concepts are lexicalized equivalence classes that (a) can be communicated and (b) shared by a group of individuals. Examples vary from *chairs* (obviously, chairs are nonidentical but merely equivalent in some way) to *odd numbers* to extremely abstract concepts, such as *cause* or *effect*.

Given the importance of concepts to human intellectual life, it is necessary to ask: Where does the ability to use concepts start, and how does it develop? It is also important to ask about the role of words in this process. Do words help us to carve up the world and form general categories? Do we apply words to already formed general categories? Or do both processes coexist? This chapter attempts to answer some of these questions.

Writing of this chapter is supported by IES Grant R305A140214 and NIH grants R01HD078545 and P01HD080679 to Vladimir Sloutsky. I thank members of the Cognitive Development Lab for helpful comments.

Principles of Conceptual Development

I begin by providing five principles that guide my review of conceptual development.

1. There is diversity of conceptual behaviors that range from relatively simple and universal to complex and uniquely human. Because this chapter is based on the assumption that simpler forms are a foundation for more complex forms, I review multiple forms of conceptual behavior.
2. Simpler forms (such as generalization) are more universal than more complex forms, and they exhibit early onset in the course of individual development. More complex forms (such as conceptual hierarchies of lexicalized categories) are unique to humans. They exhibit late onset in the course of ontogenesis and are likely to depend on other aspects of cognitive development, including the development of attention and memory.
3. The development of more complex forms of conceptual behavior is more likely to be affected by language and instruction than is the development of simpler forms.
4. The structure of input matters: Learning of statistically denser categories exhibits early onset, is present in a broad variety of species, and does not require instruction. In contrast, learning of more statistically sparse categories exhibits later onset, may be limited to organisms with functioning prefrontal cortices, and may require instruction.
5. Conceptual development progresses from less structured representations of concepts (i.e., similarity-based representation of concepts that have few links to other concepts) to more structured representations of concepts (i.e., these may involve hierarchies, taxonomies, and other complex structures).

Structure and Organization of This Chapter

Having laid out the theoretical principles that guide this chapter, I turn now to a brief preview of the chapter. I begin with an overview of theoretical approaches to concepts. I then turn to a discussion of multiplicity of conceptual behavior and to conceptual development during infancy and postinfancy childhood. When reviewing conceptual development, I consider the role of language in conceptual development, in the acquisition of semantic knowledge and of conceptual hierarchies, and in the use of concepts in reasoning.

THEORETICAL APPROACHES TO CONCEPTS AND THEIR DEVELOPMENT

The study of concepts and their development has a long history (see Goldstone, Kersten, and Carvalho, 2017 for the most recent review). Similar to many other fields of study, the study of concepts originated in philosophy, and for a long time it remained a purely philosophical endeavor. Therefore, early theories of conceptual development were strongly influenced by ideas that originated in philosophy.

Early Psychological Theories

Classical Approach: Piaget and Vygotsky

The main philosophical idea that influenced early theories of conceptual development is the logic of classes, a foundation of syllogistic reasoning. According to the logic of classes, classes of progressively increasing generality can be created by means of abstraction. For example, a boy (defined as a young human male) can be included in a more general class of human males, which in turn can be included in a more general class of humans, and so forth. Therefore, membership in a more specific class is defined by a combination of a superordinate class (e.g., being a human male in the case of a boy) and a distinctive feature (e.g., being young). Each is necessary, and the two are jointly sufficient to determine membership in a class. These necessary and sufficient features form a *definition* of a concept. If concepts are based on definition, then the goal of conceptual development is to discover these definitions. The logic of classes presumes at least three additional organizing principles. The first is class inclusion: Subordinate classes can be properly included in superordinate classes (e.g., all children are people). Second, any more general (or superordinate) class consists of a finite number of more specific (or subordinate) classes that are exhaustive of this general class. For example, the superordinate class of humans can be broken down into women and men that fully exhaust the class of humans. And finally, the subclasses of a larger class are mutually exclusive—they do not have common members. It is easy to notice that common quantifiers, such as *all*, *some*, and *none*, express all these relations. For example, the term "all" expresses the relation of a subordinate class to a superordinate class (e.g., "All men are humans"), "some" expresses the relation of a superordinate class to a subordinate class (e.g., "Some humans are women"), and "none" expresses the relation between the two mutually exclusive classes (e.g., "None of the men are women"). If the concepts are classes and mature conceptual organization is governed by the logic of classes, then a theory of conceptual development must explain how individuals acquire the logic of classes.

Two of the most influential early theories are those of Piaget and Vygotsky. Both Piaget (e.g., Inhelder & Piaget, 1964) and Vygotsky (1934/1986) attempted to link conceptual development to acquisition of the logic of classes and to the discovery of definitions. Therefore, there is little surprise that both authors believed that development progresses from less organized to more organized logical thought, from failing to understand class inclusion and mutually exclusive relations among subsets to appreciation of these relations.

Although the idea that mature concepts are based on definitions and the logic of classes advanced the study of concepts and their development, by the mid-1970s, this "classical" approach started running into difficulties. These difficulties and ideas that eventually replaced the classical approach are reviewed in the next sections.

Subsequent Theoretical Development

Prototypes, Exemplars, and Theories

In their book, *Categories and Concepts*, Smith and Medin (1981) reviewed the status of the classical view as a theory of conceptual structure. They concluded that given a large number of problems that the classical view runs into, it cannot contend for being a theory of concepts. Although I do not fully review these difficulties here (for such a review, see Medin, 1989), I offer a quick reminder of them. First, for most everyday concepts, it was impossible to come up with a set of necessary and sufficient features shared by all examples of the concept. Second, contrary to the classical view that all examples would be equally good instantiations of a concept (because all possess the defining features of the concept), observations showed that people treat examples differently: They may consider an apple to be a better example of fruit than a kiwi. And third, there are unclear cases (e.g., is floor lamp furniture or appliance?), which should not exist if concepts are organized in accord with the classical view. For example, should a rug be considered furniture? Is a rotten egg still food? These and other problems led researchers to consider alternatives to the classical view. Two are the probabilistic and the theory positions, each considered next in turn.

Probabilistic Approach: Prototypes and Exemplars

As summarized by Medin (1989), the probabilistic view holds that many categories are ill-defined, which means that there is no clear-cut category-inclusion rule but rather features are probabilistically distributed within and across categories (hence the name "probabilistic"). In the absence of a defining feature (i.e., a feature shared by all members of the category but by none of the nonmembers), categories are organized according to family resemblance, which means that each shared feature is common to many but not to all members of the category.

If there are no defining features, how are categories learned? According to this view, categories are clusters of correlated attributes, and people are capable of detecting these clusters (Rosch & Mervis, 1975). Although researchers working within the probabilistic approach generally adhere to these ideas, they vary in their proposals about how category representations are formed. Some believe that people form a summary representation of a category, which has been referred to as the *prototype*. The prototype can be the central tendency among the category members, the single best example, or the ideal instance that possesses all of the characteristic features of a category. The prototype plays a critical role in categorization decisions: If a candidate item is similar enough to the prototype, it is classified as the member of a category (J. D. Smith & Minda, 1998, 2000).

Another way of conceptualizing probabilistic categories is the exemplar view (e.g., Medin & Schaffer, 1978; Nosofsky, 1986). According to this view, no summary representation is formed and participants keep a memory record of all encountered members of a category, or category *exemplars*. If a new item is seen to be more similar to stored exemplars of the category than to stored nonexemplars, the item is judged to be a member of a category.

These two approaches have complementary strengths and weaknesses, and there is considerable literature comparing the prototype and the exemplar approaches. (See Wills & Pothos, 2012, for a recent review.) Given that the differences between the two approaches are rather small (especially when both are compared to the other approaches), I will not focus on these differences here. At the same time, it is worth mentioning that some researchers (Murphy & Medin, 1985) have criticized the very principle that gives rise to both approaches. They offered instead an alternative known as the knowledge-based approach to concepts.

Knowledge-Based Approach: Concepts Are Organized by Theories

Medin (1989) expressed what is perhaps the most central idea of this approach: "Classification is not simply based on a direct matching of properties of the concept with those in the example, but rather requires that the example have the right 'explanatory relationship' to the theory organizing the concept" (p. 1474). Therefore, people may pay attention to clusters of correlated features not because features are correlated but because correlations suggest that there is an underlying cause responsible for these correlations, and people may believe that it is this cause that is the central (or essential) feature that determines the membership in a category. For example, wings, hollow bones, and the ability to fly correlate, as do fins, gills, and the ability to swim. These correlations may suggest that there is *something* (although they may not know what it is) that gives rise to these correlations. In addition, because people have knowledge or intuition about how different kinds of categories (e.g., natural kinds or artifacts) are organized, they may assume that radically different kinds of features are central for different kinds of categories.

Based on these ideas, some (e.g., Gelman, 2003, 2004) suggested that even young children hold "theoretical assumptions" that drive their learning of categories. These assumptions are likely to be a priori in that they are preconditions rather than consequences of learning.

This approach to concepts presumes that both acquisition and use of even simple categories requires much background knowledge. Although this knowledge-based approach is highly appealing and has left a large footprint in the study of conceptual development, it is not uncontroversial. One frequent criticism is that it uses complex conceptual knowledge (i.e., the mentioned "theoretical assumptions") that itself needs an explanation as an explanatory primitive. (See Sloutsky, 2010; L. B. Smith & Heise, 1992; Spencer et al., 2009.)

Summary

The early theories of concepts assumed that concepts are based on the logic of classes and have necessary and sufficient features. These features were believed to define a concept and distinguish it from the other concepts. Conceptual development was considered a process of acquisition of the logic of classes and of organizing the concepts according to this logic. However, additional work suggested that concepts may not be organized this way: People have many concepts that do not have defining features (or at least experts fail to find them). The demise of the "classical view" of concepts led to two alternative arguments. Some argued that concepts are clusters of correlated features and that they are organized probabilistically. Others have argued that people interpret feature clusters as caused by deeper features, and they believe that these deeper causal features determine category membership. However, whatever position is taken, it remains necessary to explain conceptual development. When do concepts emerge? How do they change? What is it that develops? These are topics of subsequent sections. The next section reviews the multiple manifestations of conceptual behavior.

MULTIPLE MANIFESTATIONS OF CONCEPTUAL BEHAVIOR

Conceptual behaviors come in various forms: They range from more simple, universal, and early-emerging forms (i.e., establishing equivalence between nonidentical percepts) to rather complex, uniquely human, and late-emerging forms (i.e., forming a conceptual network in a knowledge domain). Although the multiplicity of conceptual behaviors is widely accepted, the relationships between simpler and more complex forms are not well understood. Do more complex forms emerge from simpler forms, or are these forms independent? The goal of this section is to capture this broad range of conceptual behaviors and to consider answers to the question of relationships between simpler and more complex forms of conceptual behavior.

Category Learning and Category Knowledge

Is there any commonality between perceptual groupings that are based on luminance (e.g., Figure 2.1A) and young children's intuitions about whether animals and plants are alive?

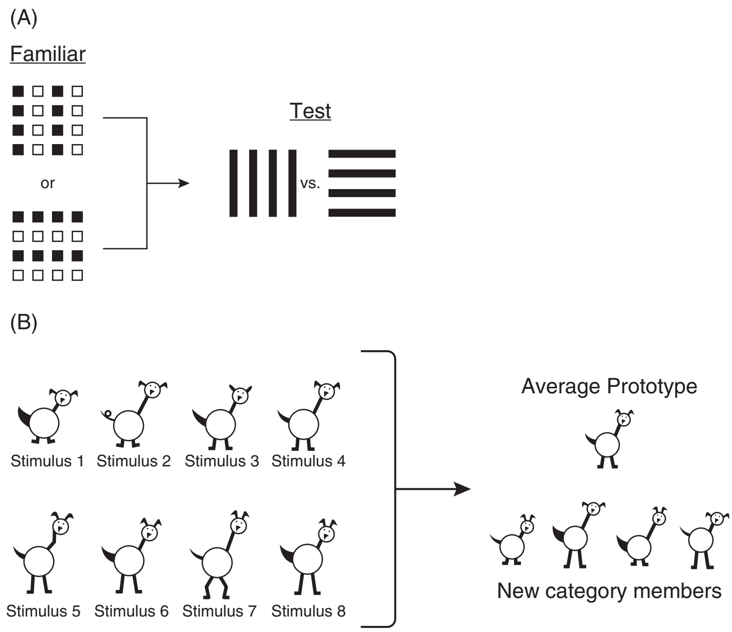


Figure 2.1 A. Example of luminance-based categories used in Quinn and Bhatt's (2006) study. Participants were familiarized with either column-based or row-based organization of squares and then tested on new stimuli that preserve the trained luminance-based organization. B. Creature-like categories used by Younger (1990). Participants were trained on a set of stimuli on the left and tested with stimuli on the right.

I suggest that studies investigating learning of perceptual groupings and those investigating naive beliefs deal with different aspects of the same problem. The former studies try to understand how people acquire new categories, whereas the latter try to understand how people use and deploy existing concepts and conceptual networks in their thinking about the world.

Therefore, an important distinction to consider is between learning new categories and using existing categories. For example, a person may learn *de novo* that eagles and hawks are two different categories of birds or participants may come to a study equipped with this distinction and merely deploy their knowledge when categorizing large birds. Category knowledge is informative with respect to what people know, whereas category learning is informative with respect to what can be learned, how, and when. These types of conceptual behavior prompt different developmental questions. Category learning prompts questions of how people acquire, store, and use categories across development and whether the mechanisms of category learning change with development or remain the same. Category knowledge prompts questions of what children of different ages know in different knowledge domains, how this knowledge is organized, and what the sources of this knowledge are. Therefore, studies of how people use categories versus studies of how people learn categories stem from somewhat different research traditions. Although both issues are important, the latter issue is a more basic one: Even if one studies existing knowledge in a particular knowledge-rich domain such as naive biology or naive physics, the question of how people acquired that knowledge in the first place needs to be answered.

Perceptual Groupings, Categories, Concepts, and Conceptual Networks

It is important to note that conceptual behaviors vary in levels of complexity ranging from simple perceptual groupings to arbitrary categories, to full-blown lexicalized concepts that are linked to other concepts that thereby form conceptual networks. The study of each type of conceptual behavior requires somewhat different research paradigms.

First, people can learn perceptual groupings or equivalence classes that are based on purely perceptual properties. Such groupings may include imposing categorical boundaries on sensory continua (known as categorical perception, e.g., Eimas, 1994), learning dot patterns coming from a single prototype and generalizing learning to distortions from the studied prototype, or forming a category based on image properties. (See Bhatt & Quinn, 2010, for a review.) Such groupings typically are studied using the *generalization* paradigm, which is perhaps the simplest conceptual task. In this task, a participant first learns a single grouping (i.e., category A) and then decides whether a new stimulus is a member of A or not. Therefore, such paradigm is sometimes referred to as A versus non-A task. For example, a participant can be familiarized with cats and tested on cats versus dogs. As I discuss later, most infant studies examining category learning use this kind of task. This is the simplest form of categorization because it is possible to extend category membership on the basis of global familiarity. At the same time, it is difficult to know within this paradigm how many features control categorization responses and what these features are.

A more complicated variant of conceptual behavior requires one to learn two or more mutually exclusive categories (e.g., cats versus dogs) at the same time. This task often is referred to as A versus B categorization. The categories are mutually exclusive because there are no members common to A and B (i.e., $A \cap B = \varnothing$). This task is more difficult than A versus non-A because a decision of whether a novel item belongs to A or to B cannot be made on the basis of global familiarity (i.e., both A and of B are equally familiar). The studied categories can be based on multiple correlated features (birds have wings, feathers, and beaks, and fish have scales, fins, and gills), few features (e.g., squirrels have a long, fluffy tail, and hamsters have a small tail), or relations among features (e.g., rectangles can be grouped into tall if the aspect ratio is less than 1 and wide if the aspect ratio is more than 1). The categories also may be deterministic (such that there is a subset of features that is sufficient to predict category membership with a 100% accuracy) or probabilistic (such that any feature or a combination of features predicts category membership only with a degree of probability). Therefore, to make a categorization decision, at the very minimum, some processing of two category structures is required. This task has been used in some studies with infants and animals and in many category-learning studies with children and adults.

An even more complicated variant of conceptual behavior is the ability to lexicalize categories and use them in reasoning, inference, prediction, or judgment. Such lexicalized categories can be defined as concepts proper. Lexicalization is critical as it enables acquiring knowledge that may not be directly observable in a given situation (e.g., dogs are friendly pets, they like meat, and they are taken to a vet for a physical exam). In other words, having a word for a category allows accumulation of knowledge from sources that are not based on direct observation of category members. These sources include conversations with others, reading, and formal education. Such concepts proper can be studied in a variety of tasks, including grouping of items, property listing, picture naming, and many others. A grouping task may require participants to put together items of the same kind (e.g., toys versus animals), whereas an attribute listing task may require a participant to list properties of categories (e.g., of cats, birds, or animals).

Finally, a conceptual network involves not only knowledge of concepts but also of relations among these concepts. Take, for example, Newton's second law ($F = ma$) that acceleration of a body is directly proportional to the net force acting on the body and inversely proportional to the mass of the body. Here the concepts of mass, force, and acceleration are linked together in a conceptual network. Such networks can be organized in a variety of ways; for example, networks of naturally occurring categories often have hierarchical, or taxonomical, organization (e.g., greyhound → dog → mammal → animal → living thing). One way of detecting such hierarchies is a classification task in which a diverse set of items is partitioned into *n* mutually exclusive and exhaustive subsets. These subsets can then be further partitioned into smaller groups or combined into larger groups.

A conceptual hierarchy is a variant of an advanced conceptual organization, and it depends critically on mastering the relation of class inclusion. The term "class inclusion" refers to a situation when a subset of items (s_1) is properly included in a larger set (S) so that ($s_1 \leq S$), as in German shepherds are dogs, and the mastery of class inclusion is examined in class inclusion tasks. Conceptual hierarchies are related to reasoning with quantifiers (i.e., *All* members of s_1 are members of S, but *some* members of S *are not* members of s_1). However, it is not known whether the mastering of class-inclusion relations is necessary for understanding of the meaning of quantifiers *some*, *all*, *none*, and *at least one*—something that Piaget believed (Inhelder & Piaget, 1964)—or, alternatively, whether acquisition of quantifiers bootstraps the development of class-inclusion relations necessary for forming conceptual hierarchies. It also has been argued that classification tasks may underestimate children's concepts: The fact that a child may put together a dog and a bone does not mean that the child considers the two to be the same thing (e.g., Fodor, 1972). However, classification tasks are useful in that they may reveal a limit on the kinds of concepts children may form.

Although it is tempting to consider perceptual groupings, categories, concepts, and conceptual networks as qualitatively different conceptual behaviors, this chapter argues that this is not the case and that there is continuity among these instantiations of conceptual behavior. According to this view, human concepts develop from perceptual groupings (something that also can be achieved by certain nonmammalian species) to conceptual networks that are likely to be unique to humans. One important goal of this chapter is to elucidate such development.

Different Kinds of Categories

Are all categories the same? The standard answer to this question is yes. Here is an example of this point expressed by Shipley (1993): "Three psychological properties appear to characterize categories: (1) they have labels that are used to identify objects, (2) they serve as the range of inductive inferences, and (3) their members are believed to share a 'deep' resemblance" (p. 266). However, nonhuman animals and prelinguistic infants can form perceptual categories (Lazareva & Wasserman, 2008; Quinn, 2002a). This fact suggests that labels are not a necessary component of categories. In addition, people (as well as nonhuman animals) can learn arbitrary memory-based categories (e.g., items in their living rooms), suggesting that deep resemblance is not necessary either. Therefore, the kinds of categories that people can and do learn is quite broad, and it may include different kinds.

Although there is little doubt that categories differ in content, the most interesting distinctions pertain to category structure. Structural differences identified by researchers include syntactic differences (nouns versus verbs; e.g., Gentner, 1981), ontological differences (natural kinds versus nominal kinds; e.g., Kripke, 1972), taxonomic differences (i.e., basic level versus superordinate level; e.g., Rosch & Mervis, 1975), differences in organizational principle (entity categories versus relational categories; e.g., Gentner & Kurtz, 2005), differences in concreteness (concrete versus abstract categories; e.g., Barsalou, 1999), differences in category coherence and confusability (e.g., Homa, Rhoades, & Chambliss, 1979; Rouder & Ratcliff, 2004; Smith & Minda, 2000), and some other distinctions.

Kloos and Sloutsky (2008) proposed another structural distinction, one that could form the basis for many of the preceding distinctions. They proposed the idea of statistical density, that is, a measure of category structure that (a) can (in principle) be measured independently rather than be inferred from participants' patterns of response and (b) provides a continuous measure rather than a dichotomous one (which makes it well suited for capturing the graded nature of differences between categories).

Conceptually, statistical density is a ratio of variance relevant for category membership to the total variance across members and nonmembers of the category. Intuitively, statistical density is a measure of how members of a category are separated from nonmembers. A brief overview of statistical density ways of calculating it is presented next; a more detailed discussion is presented elsewhere (Kloos & Sloutsky, 2008). Three aspects of stimuli are important for calculating statistical density: variation in stimulus dimensions, variation in relations among dimensions, and attentional weights of stimulus dimensions.

First, a stimulus dimension may vary either within a category (e.g., members of a target category are either black or white) or between categories (e.g., all members of a target category are black, and all members of a contrasting category are white). Within-category variance decreases density; between-category variance increases density.

Second, dimensions of variation may be related (e.g., all items are black circles), or they may vary independently of each other (e.g., items can be black circles, black squares, white circles, or white squares). Covarying dimensions result in smaller variability (and thus in greater density) than dimensions that vary independently.

The third aspect is the attentional-weight parameter. Without this parameter, it would be impossible to account for learning of some categories. In particular, when a category is dense (i.e., when multiple dimensions are correlated within a category), even relatively small attentional weights of individual dimensions add up across many dimensions. This makes it possible to learn the category without supervision and without attention to a particular dimension. Conversely, when a category is sparse, only a few dimensions are relevant (i.e., members of a category are all red but vary on multiple dimensions, such as shape, color, texture, and size). If attentional weights of each dimension are too small, some guidance (or *supervision*) could be needed to direct attention to the relevant dimensions.

The idea of statistical density has important implications for the development of category learning. One possibility is that category learning progresses from spontaneous learning of highly dense categories to less spontaneous (and more guided or supervised) learning of more sparse categories.

Spontaneous versus Supervised Category Learning

Theoretically, category learning is considered *supervised* when (a) categories are marked or labeled and (b) participants are given feedback when they assign items to categories. In contrast, category learning is considered *unsupervised* when participants are only presented with items, without classes being labeled or feedback being provided.

The idea (supported by evidence) that dense categories can be learned without supervision has an important implication: Prelinguistic infants should be able to implicitly learn many categories by interacting with the world surrounding them. Incidentally, the very first nouns that infants learn denote these dense categories (see Dale & Fenson, 1996; Nelson, 1974). Therefore, it is quite possible that some early word learning consists of learning lexical entries for already known dense categories.

At the same time, many other concepts that are based on sparse categories (these include multiple legal, ethical, mathematical, and scientific concepts) are unlikely to be learned spontaneously. Learning of these concepts requires various degrees of supervision, and it is likely that many of these concepts are learned in the course of formal schooling. An interesting case is a set of naive scientific concepts (e.g., naive biology), which are naive conceptual networks in domains studied by science (e.g., Hatano & Inagaki, 1994). Although there is little doubt that even preschoolers have some of these naive concepts (e.g., the concept of a living thing), the origin of these concepts is not well understood. Are these concepts acquired spontaneously through experience with various kinds of plants and animals? Or are these concepts learned in a supervised manner, with supervision being offered by parents, children's books, television, and perhaps some other sources? Currently we do not have definitive answers to these questions, but it seems highly unlikely that categories of such low statistical density are acquired spontaneously, without supervision (cf. Opfer & Siegler, 2004).

Summary

This section reviewed the multiplicity of conceptual behaviors. It considered distinctions (1) between category learning and category use, (2) among different types of conceptual behaviors (e.g., perceptual groupings, categories and concepts), and (3) among different kinds of category structures as well as the ways these structures can be learned. In the sections to follow, I review neural mechanisms of categorization, basic categorization abilities in nonhuman species, category learning in infancy, and lexical and semantic development.

CONCEPTUAL DEVELOPMENT IN INFANCY

If forced to reduce conceptual behaviors in infancy to two primary findings, I would list the ability of preverbal infants to learn categories at all and their ability to learn many of these categories without a teaching (or supervisory) signal. These and other issues are discussed in the next sections.

Preverbal Infants Exhibit Evidence of Category Learning

How do we examine conceptual behaviors in a nonlinguistic organism that cooperates for only a short period of time? This is not an easy task. Category learning in human infants typically is examined using a wide range of stimuli and a wide range of research paradigms. Stimuli used with infants typically make use of sensory-defined categories, pictures of animal-like creatures (e.g., see Figure 2.1B), and real objects. Research methods include visual attention, object examination, sequential touching, and operant conditioning paradigms. In visual attention paradigms, infants are first familiarized with (or habituated to) members of a to-be-learned category. They are then presented with either a novel item from the studied category or an item from a nonstudied category. Learning is inferred if the infant displays (a) longer looking to an item from a nonstudied category coupled with (b) the ability to discriminate familiar from novel members of the studied category. (See Chapter 1 in this volume for a broad overview of visual attention.) Object examination and sequential touching paradigms (Rakison & Butterworth, 1998; see Cohen & Cashon, 2006, for a review) are based on a similar logic, but participants are presented with toy replicas of objects and given an opportunity to examine these replicas. "Examining" often is defined as focused looking in the presence or absence of manipulation (Cohen & Cashon, 2006).

In the sequential touching paradigm, the infant is presented with replicas of objects from two categories (e.g., horses and cows) and is given an opportunity to examine these objects. The sequential order in which the infant examines the objects serves as the dependent variable. Any deviation from randomness (i.e., greater probability of examining objects within a category than across categories) is taken as evidence that the infant is responding on the basis of the category.

Two important findings stem from research using these paradigms. First, pioneering studies using visual attention paradigms by Eimas, Cohen, and their colleagues have shown that, at least by about 10 months of age (and often as early as by 3 months of age), infants can learn various animal categories (Eimas & Quinn, 1994; Quinn, Eimas, & Rosenkrantz, 1993; Oakes, Coppage, & Dingle., 1997) as well as more artificial categories of patterns of luminance (Bhatt & Quinn, 2010, for a review), geometric shapes (Bomba & Siqueland, 1983; Quinn, 1987), schematic animals (Younger, 1990; Younger & Cohen, 1985), and schematic faces (Strauss, 1979).

And second, *what* babies learn depends on the input. For example, Quinn and colleagues (e.g., Quinn et al., 1993) found important asymmetries in category learning. (Perhaps the most striking one is that babies often learn categories of cats that exclude dogs but not of dogs that exclude cats.) This finding is important because it clearly indicates that infants are sensitive to category structure. Although making a significant step of demonstrating infants' sensitivity to structure, the study did not reveal which aspects of the structure infants are sensitive to. Subsequent work (French, Mareschal, Mermillod, & Quinn, 2004) provided answers to this question. French et al. (2004) found that *what* infants learn is affected by variability in the input: Greater featural variability in the input is accompanied by learning of a broader category. In particular, cats' features tend to vary less than those of dogs. As a result, when presented with cats, babies tended to learn a narrow category of cats; when presented with dogs, babies tended to learn a broader category of cats-and-dogs. These findings were corroborated by subsequent experimental and computational work in which researchers created artificial sets of broadly varying cats and narrowly varying dogs, resulting in a reversal of the asymmetry (French et al., 2004).

Infants' sensitivity to input variability also was seen in an object examination task.

For example, Oakes, Coppage, and Dingel (1997) found that 10-month-olds were more likely to dishabituate to a novel out-of-category item if the set of items used in the study was uniform than when it was variable. However, it was not clear from this study whether the more variable input resulted in a failure to learn the target category or in learning a broader, more inclusive category.

Although variability of input is important, it is not the only factor that affects infant category learning. For example, Gliozzi, Mayor, Hu, and Plunkett (2009) and Mather and Plunkett (2011) demonstrated both computationally and experimentally that even with the same set of items, the order in which items are presented affects what infants learn about a category.

In sum, infants are highly sensitive to the structure of input. In addition, as argued by Quinn (2002b), infants can do more than just detect and discriminate; they also can group, organize, relate, and generalize stimuli in their environment and form perceptual categories. Is this evidence of conceptual behavior? I believe that yes, perceptual categories are the starting point of conceptual development. However, as I discuss later, this conclusion is not uncontroversial: Some researchers believe that perceptual category learning has little, if anything, to do with conceptual development.

Preverbal Infants Can Learn Categories Without Teaching or Supervisory Signal

As discussed, one of the central findings of infancy research of the past 20 years is that infants can learn categories without a teaching (or supervisory) signal. Supervised and unsupervised learning may result in different representations of a category in neural networks (e.g., Japkowicz, 2001) and in human learners (e.g., Kloos & Sloutsky, 2008). Most infancy studies use unsupervised learning: Infants generally are familiarized with category exemplars and then tested either on a new member of the studied category or on a novel item. The fact that infants exhibit preference for a novel item indicates that they can learn a category without supervision. It should be noted, however, that most studies demonstrating the ability of infants to learn categories familiarized infants with only a single category. Despite its many advantages, this paradigm has a number of limitations. Most important, category learning is inferred from a preference for a novel item, and, therefore, much depends on the choice of the novel item. In many situations, it is difficult to interpret what exactly was learned. For example, consider an experiment in which 10-month-olds are familiarized with balls varying in color and size and tested on balls versus flowers. Further suppose that participants exhibit reliable novelty preference, looking longer at a flower than at a new ball. Although it is clear that participants learned something, what exactly they learned is less clear. Is it the category of balls, round things, things without parts, or things with uniform texture? Similar problems arise when infants learn natural kind categories.

One way of addressing this problem is to present infants with the task of learning two categories simultaneously. Although this approach has challenges and still is rarely used in the study of infant categorization, the few existing studies using this method have been encouraging. In one study, Plunkett, Hu, and Cohen (2008, Experiment 2) presented 10-month-olds with a stimulus set consisting of two categories. (See Figure 2.2 for an example: The four items on the left are members of Category A, and the four items on the right are members of Category B.) Because items had continuous dimensions (e.g., neck length or ear separation), a test item could be either an extreme case of Category A (and thus far from Category B) or fall

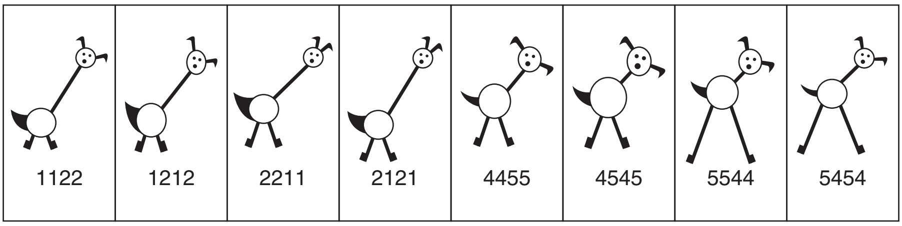


Figure 2.2 Stimuli used in Plunkett et al. (2008).

between categories A and B. When presented with these item types, infants exhibited a preference for the in-between category items, thus suggesting that they had learned two categories.

The second example comes from a study by Sloutsky and Robinson (2013), who used a variant of "switch" task (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). These researchers presented 14-month-olds with two categories, one defined by the same color and another defined by the same shape. Here the two categories were presented in different contexts: Items from Category A were presented on one background, in a certain location on the screen and with a particular kind of ornamentation (border) around them; items from Category B were presented on a different background, in a different location on the screen and with a different ornamentation. At test, participants were presented with (a) the same trials (new members of studied categories), (b) new trials (entirely new items), and (c) switch trials (new members of a studied category presented in the context of the other category). Learning was inferred when participants exhibited novelty preference on switch and new trials but not on same trials. Participants indeed exhibited this pattern, indicating that they succeeded at learning both categories.

The third example comes from a study by McMurray and Aslin (2004), who introduced a two-alternative anticipatory eye-movement paradigm. In this paradigm, one category is associated with one outcome (e.g., an engaging object appearing on one side of the screen), and another category is associated with another outcome (i.e., another engaging object appearing on another side of the screen). Category learning is inferred from anticipatory looking to the correct side of the screen when a member of one of the two categories is presented. McMurray and Aslin reported successful learning of two categories by 5- and 7-month-old infants.

Note that in none of these paradigms were participants explicitly given a teaching signal or explicitly reinforced for a correct response. Therefore, taken together, these findings present strong evidence that a teaching signal is not necessary for category learning in infancy. These findings raise another important question: To what extent can infants benefit from supervision? I address this question in the section on the role of language in infant category learning.

Controversial Issues in Infant Category Learning

Although the question of whether infants can learn categories is relatively uncontroversial—they do!—questions pertaining to how infants learn categories and how these categories relate to later conceptual development have generated considerable disagreement. These points of disagreement pertain to (1) the way infants learn and represent global categories, (2) whether category learning in infancy is a continuous or a discontinuous process, and (3) the role of language in infant category learning. Only some of these issues have been resolved to date.

Nature of Global Categories in Infancy

There exists a large body of evidence that young infants can form basic-level categories, such as cats or dogs, and more global-level superordinate categories, such as animals or vehicles. According to one view (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), there is a developmental progression from mastering basic-level categories (e.g., cat or truck) to superordinate categories (e.g., animal or vehicle).

According to another view, the progression is in the opposite direction: More global superordinate categories are acquired prior to basic-level ones (Mandler & Bauer, 1988). Although this idea is reasonable, there is evidence that much younger infants (sometimes as young as 3 months of age) can learn basic-level categories drawn from the same superordinate category, such as cats versus dogs (Quinn et al., 1993). However, notice that Mandler and Bauer used manual exploration procedures, whereas Quinn et al. (1993) used visual attention procedures. It is possible therefore that basic-level categories can be learned perceptually while superordinate-level categories cannot (because the latter have too much perceptual variability to be picked up by the perceptual system). Hence, depending on the learning procedure, infants may learn different types of categories—"perceptual" categories in the course of visual exploration and "conceptual" categories in the course of manual exploration. Although this possibility is not unreasonable, currently it cannot account for a number of important findings.

Perhaps the most critical findings are that very young infants can, in fact, learn global-level categories by means of perception. For example, Behl-Chadha (1996) presented 3- to 4-month olds with a variant of the visual familiarization task and found that infants successfully formed a global category of mammals that included novel mammals but excluded other nonmammalian animals, such as birds and fish. Quinn and Johnson (2000) reported similar findings for 2-month-old infants. Critically, not only were young infants able to learn these global categories, but their ability to learn these categories appeared to come online before the ability to learn constituent basic-level categories. When Quinn and Johnson (2000) modeled these data using an auto-associator network (a simple network that learns to output the input or its part), the network also learned global categories before learning basic-level categories. (See also Rogers & McClelland, 2004.) These findings are important because the network had only perceptual input and yet was capable of learning global-level categories before learning basic-level categories. Taken together, results reviewed in this section strongly suggest that perceptual information in global-level categories is sufficient to allow very young infants and networks to learn these categories by perceptual means.

Continuity versus Discontinuity (or Monism versus Dualism) in Infant Category Learning

The fact that infants can learn both basic-level and global-level categories generated another controversy. Some researchers (e.g., Mandler, 1992) suggested that categories learned by very young infants are perceptual in nature, and categories of older infants, children, and adults are conceptual in nature (i.e., are based on more abstract, nonperceptual features). According to this account, the latter categories have very little in common with the former categories. Other researchers (e.g., Eimas, 1994; see also Quinn, 2011) rejected such dualism, suggesting instead that conceptual categories develop out of perceptual categories. As to be discussed, substantial evidence is generated by each account.

For example, Mandler (1992, 1999) offered her account of category learning. (But see Müller & Overton, 1998, for a review and critique of this approach.) The central idea of this proposal is that true concepts cannot emerge from perceptual categories and must have conceptual primitives as their starting point. These conceptual primitives are a result of perceptual analysis, which is "a process in which a given perceptual array is attentively analyzed, and a new kind of information is abstracted. The information is new in the sense that a piece of perceptual information is recoded into a non-perceptual form that represents a meaning" (Mandler, 1992, p. 589). Representations that result from perceptual analysis are called "image schemas." These image schemas (e.g., SELF-MOTION, ANIMATE-MOTION, CAUSED-MOTION) can be derived from perceptual structure but cannot be reduced to it. In turn, concepts, such as animacy, inanimacy, or agency, are built from these conceptual primitives. Evidence supporting these ideas comes from a set of studies conducted by Mandler and her colleagues (Mandler & McDonough, 1996; 1998; McDonough & Mandler, 1998) in which 11- to 14-month-old infants generalized properties (e.g., drinking) to a broad category, such as animals. Under the assumption that there is very little perceptual commonality among members of these global categories, it was concluded that these generalizations could be made only on the basis of conceptual information.

In opposition to this approach, Eimas (1994) offered a view according to which conceptual knowledge has its origins in perception. First, as discussed, very young infants can acquire both basic-level and more global categories of natural kinds by perceptual means, and it is possible that development consists of quantitative enrichment, not a qualitative transformation of these early categorical representations. Second, in principle, perceptual and associative processes can result in more abstract representations. For example, biological motion (which is clearly a perceptual category) may form the basis for a representation for animate beings. In other words, perceptual categories acquired very early in development may give rise to more abstract categories acquired later in development. In sum, according to this view, conceptual knowledge may evolve from perceptual origins if development is considered a sequence of events rather than a two-step process. Although the controversy remains unresolved, each side of the debate has generated interesting research in support of its position.

Role of Language in Infant Category Learning

The third controversial issue is the role of language in early category learning. The issue is of critical importance because it has implications for understanding the role of language in cognitive development, the nature of early category learning, and the extent to which supervision may affect early category learning. Given that this controversy is not unique for conceptual development in infancy, I return to it again later when reviewing conceptual development after infancy.

Some researchers suggest that from early in development, words are "names" of objects and categories (Balaban & Waxman, 1997; Waxman & Booth, 2003; Xu, 2002). At the computational level (Marr, 1982), this approach assumes that words function as supervisory signals directing and guiding learning. Thus, if two discriminable items share the same count noun (e.g., both are called "a dax"), the name serves as a top-down signal that the items are equivalent in some way (cf. Gliga, Volein, & Csibra, 2010). Similarly, if two items are labeled differently (e.g., "a dax" versus "a fep"), the names serve as a top-down signal that the items are different.

Another possibility is that early in development, words, just like any other perceptual feature, are first and foremost part of the input, and they influence categorization in a bottom-up, nonsupervisory fashion (Colunga & Smith, 2005; Plunkett et al., 2008; Sloutsky & Fisher, 2004a). Under some conditions, linguistic input may facilitate learning (Colunga & Smith, 2005; Plunkett et al., 2008; Samuelson & Smith, 1998, 1999), but under other conditions, it may hinder learning (Robinson & Sloutsky, 2007a, 2007b; Plunkett et al., 2008; Sloutsky & Robinson, 2008). According to this view, even if words start out as part of the stimulus input, they eventually may become supervisory signals (Casasola & Bhagwat, 2007; Casasola, Bhagwat, & Burke, 2009; Gliozzi et al., 2009; Mayor & Plunkett, 2010; Sloutsky, 2010; L. B. Smith & Yu, 2008).

Each of these possibilities presumes a distinct mechanism and neural architecture and, most likely, a different trajectory of development. Distinguishing among these possibilities and understanding the mechanisms underlying the effect of words on category learning is of critical importance for understanding cognitive development.

Words Are Supervisory Signals Facilitating Category Learning. One hypothesis is that words are invitations to form categories; that is, words function as top-down supervisory signals facilitating category learning. Evidence for this hypothesis comes from studies that use a variety of visual attention and object examination paradigms. Waxman and Markow's (1995) study was one of the first demonstrations of these effects. In their study, 9- to 20-month-olds were presented with a task that combined object examination and novelty preference. First participants were presented with four familiarization trials. On each familiarization trial, they were given one object to play with. During familiarization, the category structure (i.e., basic level versus superordinate) was fully crossed with labeling condition (Noun versus No Word), thus resulting in four between-subjects conditions. In one condition, all familiarization objects were drawn from a single basic-level category, such as cars, whereas in the other condition, all objects were drawn from a superordinate category that included cars and airplanes. In addition, in one condition, a label in the form of the count noun accompanied the familiarization objects (e.g., "Look, a car); in the other condition, no labels were introduced (e.g., "Look!"). Then participants were presented with a single test trial that included a new member of the familiarized category and a new member of a contrasting category (e.g., car versus airplane in the basic-level condition or truck versus lion in the superordinate condition). Results indicated that participants were above chance in all conditions, except for the Superordinate Category–No Word condition. These results led researchers to conclude that words facilitate infants' attention to superordinate categories.

Although these effects of words on category learning in infancy appear tenuous, there are two other potential sources of evidence. One of these sources has to do with putatively different effects of nouns and adjectives on categorization. In one study (Booth & Waxman, 2009), 14-month-olds and 18-month-olds were familiarized with items of the same color that were drawn either from the same basic-level category (e.g., purple horses) or from the same superordinate category (e.g., purple animals). In one condition, members of a category were referred to by a count noun (e.g., this one is a blicket), and in the other condition they were referred to by an adjective (e.g., this one is blickish). At test, participants were presented with a member of a familiar category (e.g., green horse) and a member of a novel category (e.g., purple chair). Item presentation at test was split into four time windows (i.e., 0–1 sec, 1–2 sec, 2–3 sec, and 3–4 sec). The analyses revealed greater novelty preference in the noun condition compared to the other two conditions, but only for the time window 3 (i.e., 2–3 sec after the stimulus onset). Is time window 3 special, or do the results stem from multiple comparisons?

Words Start Out as Features but Become Supervisory Signals in the Course of Development. The second hypothesis is that words start out as perceptual features affecting processing of visual input but that the effects of words may change over the course of development. Early in development, words may hinder category learning by attenuating processing of visual input; later in development, words may contribute to category learning by increasing within-category featural overlap. Critically, in both cases, words function as perceptual features. For example, Sloutsky and colleagues have presented evidence that novel labels and other sounds overshadow (i.e., attenuate) the processing of visual stimuli in young infants (Robinson & Sloutsky, 2004, 2007b, 2010; Sloutsky & Robinson, 2008). As a result, auditory stimuli (including novel words) interfere with category learning (Robinson & Sloutsky, 2007a; Sloutsky & Robinson, 2008). The overshadowing hypothesis is based on a series of familiarization and habituation studies in which infants were familiarized with compound auditory-visual stimuli (e.g., pairing a picture of a cat with a word or with a nonlinguistic sound) and were then exposed to a dishabituation stimulus that changed either the auditory or the visual component of the compound stimulus. At test, infants noticed the change in the auditory component but not the change in the visual component. Failure to dishabituate to a change in the visual stimulus when it was accompanied by a sound (but not when it was presented in silence) suggested that the auditory stimulus interfered with processing of the visual information (i.e., overshadowed it) during familiarization. It should be noted that familiar auditory stimuli, such as well-known names, do not produce such dramatic overshadowing effects in infants. Furthermore, novel words interfere with visual processing at younger ages (i.e., 10 months of age and younger), but the effect is reduced in older infants (i.e., 16 month of age and older) (Sloutsky & Robinson, 2008). Because of the increased efficiency of cross-modal processing, overshadowing weakens in the course of development (Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). For older infants and young children, overshadowing has an impact on processing of infrequent features (e.g., individual idiosyncratic visual features of category members) and not on processing of frequently recurring features (e.g., features shared by most category members). As a result, for these older participants, words may facilitate detection of what is common among category members, but they may undermine detection of individual features, thereby hindering the recognition of the distinction between familiar and new category members.

Although words may begin as features that affect the processing of visual input, they eventually may become supervisory signals. For example, Plunkett and colleagues (2008) presented experimental evidence suggesting that for preverbal infants, effects of words on category learning are not straightforward: Under some conditions, words may facilitate category learning; under other conditions, they may hinder category learning; and yet under other conditions, they do not affect category learning at all. To better understand this pattern of findings, Plunkett and colleagues (Gliozzi et al., 2009) developed a computational model to simulate these patterns of infants' responses. The model handled visual and acoustic information in an identical fashion, with no direct connections between objects and labels. In other words, the learning process was unsupervised. The pattern of novelty preferences in the simulations mimicked closely the infants' preferences. This finding suggested that an unsupervised learning device, which performs statistical computations on compound visual and acoustic stimuli, offers a viable solution to the problem of how labels influence category formation in the infant experiments. Although Gliozzi et al. (2009) provided support for the idea that words start as features, other research suggested that words do not have to remain features. As children develop, they may learn that words have high predictive power in determining a category, and, as a result, words may become supervisory signals. Although there is little disagreement among theorists that words eventually become invitations to form categories (cf. Casasola & Bhagwat, 2007; Lupyan, Rakison, & McClelland, 2007; Mayor & Plunkett, 2010; Sloutsky, 2010; Yamauchi & Markman, 1998), the precise developmental time course of this transformation remains unclear.

More recently, Deng and Sloutsky (2015b) presented evidence that the very idea that infants learn categories by extracting commonalities could be wrong. Infants were familiarized with exemplars from one category in a label-defined or motion-defined condition and then tested with prototypes from the studied category and from a novel contrast category. Eye-tracking results indicated that infants exhibited better category learning in the motion-defined condition than in the label-defined condition, and their attention was more distributed among different features when there was a dynamic visual feature compared with the label-defined condition. Furthermore, there were more gaze shifts in the motion-defined condition than in the label-defined condition. These results indicated that infants were successful in learning categories under the condition that favored more distributed (as opposed to focused) attention. This research raises important questions about the role of attention in infant category learning. These questions remain unanswered and will have to be answered in future research.

Summary

In sum, infant category learning is the first critical step in conceptual development. Category learning emerges early in life, and infants are proficient category learners. Although researchers generally agree that infants learn progressively more complex categories, many issues in the development of categorization remain a matter of debate. Among the most controversial issues are whether concepts emerge from perceptual categories learned by infants and the role of language in infant category learning.

Despite these controversies, most researchers agree that infants learn a variety of categories, some of which come to acquire conceptual significance for children and adults. Perhaps the most critical step in acquiring conceptual significance is lexicalization, or learning names for categories. These names eventually become part of category representation and "knowledge hubs" that help connect what is known about a given category. Words are also important for forming conceptual hierarchies, such as DOG → MAMMAL → ANIMAL → LIVING THINGS → OBJECTS. These conceptual hierarchies support propagation of knowledge through inductive, deductive, and transitive inference. For example, upon learning that all objects are made out of atoms, one may conclude (by deduction) that dogs are made out of atoms too. Similarly, upon learning that dogs are mammals and mammals are animals, one may conclude (by transitive inference) that dogs are animals. And finally, upon learning that dogs have white blood cells, one may infer (by induction, and thus with only a degree of certainty) that mammals have white blood cells too. It is not clear if these hierarchical relations can be expressed without language, at least without quantifiers such as *all*, *some*, and *some are not*, and I contend that language plays a critical role in conceptual development following infancy. The next section focuses on these issues.

CONCEPTUAL DEVELOPMENT AFTER INFANCY

A great deal of conceptual development takes place in postinfancy years. Obviously, there are multiple candidate sources of this development. Children continue acquiring language. They receive increasing input from multiple informal sources, including parents and other family members, peers, and books and media, to name a few. Children continue expanding their knowledge base that provides a foundation for acquisition and organization of additional knowledge. Their processing capacity (including working memory and selective attention) also undergoes substantial development. And they receive systematic input from formal educational sources, such as classroom materials and textbooks. It is likely that all these factors contribute to postinfancy conceptual development, albeit in different ways and to different degrees. In what follows, I consider the role of cognitive and linguistic factors in conceptual development, followed by a discussion of some of the specific achievements of semantic development, including the development and organization of semantic knowledge, the development of conceptual hierarchies, and the development of inductive inference.

Role of Cognitive and Linguistic Factors in Conceptual Development

One of the most striking changes in postinfancy development is a dramatic expansion of processing capacities coupled with dramatic growth of lexical and grammatical aspects of language. As was argued elsewhere (Sloutsky, 2010), these developments are likely to significantly affect conceptual development.

Role of Cognitive Factors in Conceptual Development

Young children undergo dramatic cognitive development during the postinfancy years, including the development of long-term memory (Ghetti & Lee, 2010; Newcombe, Lloyd, & Ratliff, 2007), working memory and other aspects of executive function (Carlson, 2005; Cowan, 1997), and selective attention (Hanania & Smith, 2009; Plude, Enns, & Brodeur, 1994; see also Chapter 1 in this volume). A growing body of work considers the role of selective attention and working memory in conceptual development.

Selective Attention. Given the role of selective attention in adult category learning and categorization (Nosofsky, 1986, 1988), it is likely that these developments are important contributors to conceptual development. For example, work suggests that the development of selective attention is a primary reason for differences in classification performance between 2- and 5-year-olds (Smith, 1989). This study, however, did not examine learning: Participants were presented with sets of three two-dimensional items (i.e., the items could vary on color and shape) and were asked to group together “the ones that go together.” Does selective attention matter when the task is to learn categories?

In a more recent study, Kloos and Sloutsky (2008) addressed this problem by examining the ability to learn categories of different statistical structures across development. Some of the categories had multiple overlapping features (i.e., these were statistically dense categories), and others had few category-defining features (i.e., these were statistically sparse categories). Although the researchers did not find differences between 4- to 5-year-olds and adults in learning the former categories, they did find evidence of profound differences in learning the latter categories. (See Figure 2.3.) Given that learning of sparser categories puts demands on selective attention, these findings indirectly implicate selective attention in the development of categorization. Other researchers (e.g., Hammer, Diesendruck, Weinshall, & Hochstein, 2009) reported related findings using different categories and category structures.

Another study implicating selective attention in the development of categorization (albeit with younger participants) was reported by Son, Smith, and Goldstone (2008). In this study, toddlers learned shape-based categories in one of two conditions, either through perceptually impoverished examples that communicated primarily shape information or through perceptually rich, realistic items. Participants' category learning was then tested with either impoverished or rich stimuli. Results indicated that regardless of the testing stimuli, participants exhibited more robust learning when trained with impoverished stimuli. Given that perceptually rich stimuli carry much information that is not relevant for category learning, these stimuli are likely to put greater demands on selective attention than impoverished stimuli, and young participants cannot meet these demands due to immaturity of selective attention.

However, these studies present indirect evidence for the role of attention in category learning and generalization and for developmental differences in attention. More direct evidence stems from recent work (e.g., Deng & Sloutsky, 2015a; see also Deng & Sloutsky, 2016; Rabi & Minda, 2014; Sloutsky, Deng, Fisher, & Kloos, 2015). For example, Deng and Sloutsky (2016) presented 4-year-olds, 6-year-olds, and adults with a category-learning task, in which participants learned two categories derived from

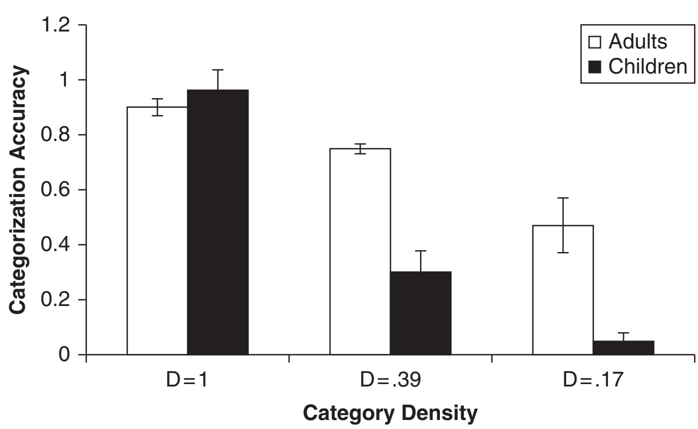


Figure 2.3 Unsupervised category learning by density and age group.
SOURCE: After Kloos & Sloutsky (2008).

two prototypes. Both categories had multiple probabilistic features (which provided only approximate information about the category) and a single deterministic feature (which provided a perfect cue to the category). Therefore, participants could learn either a probabilistic, similarity-based category that included all or many of the features or a deterministic, rule-based category that was based primarily on the deterministic feature. Category learning was probed in multiple ways, two of which are important. First, participants were tested on which category they learned. To answer this question, participants presented with *switch* items (those that have the deterministic feature of one category and probabilistic features of another). If participants learned similarity-based categories, they should respond in accordance with probabilistic features; if they learned rule-based categories, they should respond in accordance with rule-based features.

And second, participants' memory for deterministic and probabilistic features was tested. If they attended selectively to the deterministic features, their memory for these features should be better than their memory for any probabilistic feature. However, if they distributed attention across all the features, they should have comparable memory for all the features. Results (see Figure 2.4) indicate that 6-year-olds and adults learned deterministic categories while 4-year-olds learned probabilistic categories. Perhaps more important, adults and 6-year-olds remembered deterministic features better than probabilistic features, while 4-year-olds remembered all features equally well. These results present further evidence that adults (and perhaps 6-year-olds) attend selectively to deterministic features, and 4-year-olds distribute attention across multiple features.

Taken together, these results indicate that the development of attention could be a strong contributor to conceptual development. These studies have also laid groundwork for directly measuring attention across development (by using eye tracking or other related measures) and linking these measures with category learning.

Working Memory. Whereas selective attention may be important for learning visual categories, working memory could be important for (a) learning more abstract categories that are based on features that are not directly observable and (b) using these categories for inductive inference (e.g., Halford, Andrews, & Jensen, 2002; Halford, Cowan, & Andrews, 2007). For example, Halford and colleagues (2002, 2007) argued that understanding of class inclusion relations (which is necessary for both the formation of conceptual hierarchies and property induction tasks) depends on working memory capacity. Recall that class inclusion refers to a situation when a subset of items (s_1) is properly included in a larger set (S), and understanding of class inclusion requires understanding that $s_1 \leq S$ (i.e., there cannot be more German shepherds than dogs). According to this argument, property induction (If X has property P, does Y have property P?) also depends on working memory capacity, especially if "X" and "Y" are different level categories (e.g., X are dogs and Y are mammals).

The importance of working memory in induction with familiar categories also was seen in recent research by Fisher and colleagues (Fisher, Godwin, Matlen, & Unger, 2015). These researchers examined induction when familiar semantically related labels (e.g., crocodile and alligator) were provided, but appearance information was not. The primary focus was on variables that predict developmental change in this ability within individuals. The results indicated that working memory was an important predictor of the developmental change in the ability to rely on semantic information when performing

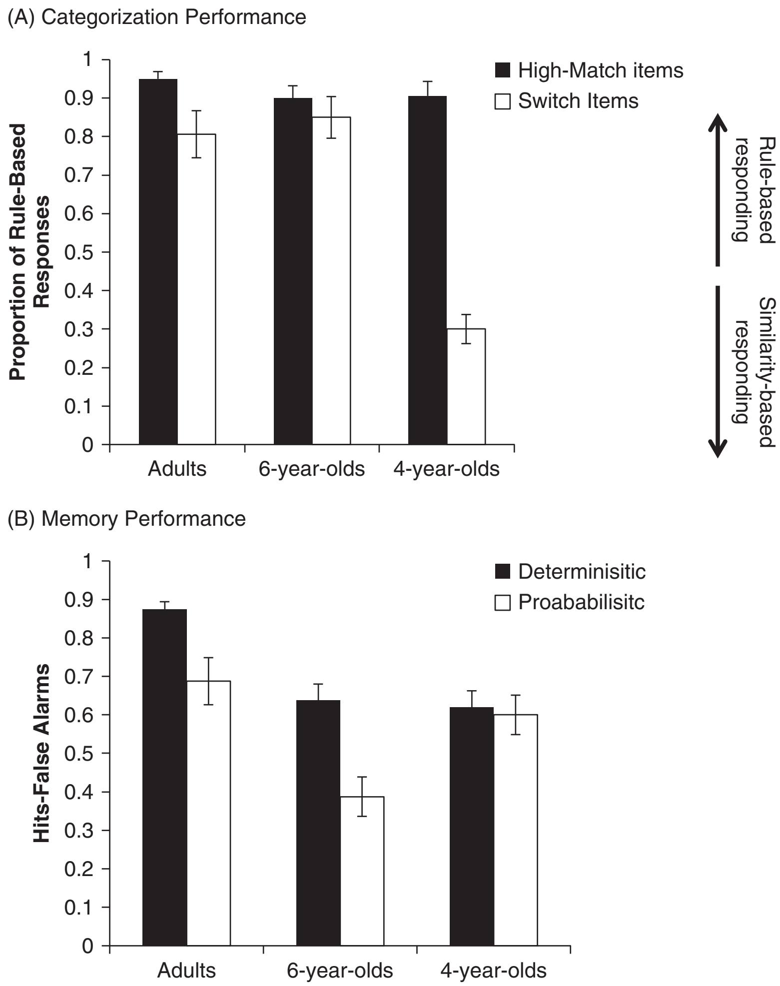


Figure 2.4 Categorization performance and memory for features across development (after Deng & Sloutsky, 2015a). A. Categorization performance: Proportion of rule-based categorization responses by trial type and age. B. Recognition performance: Memory accuracy by feature type and age.

induction. Although this evidence is indirect, it nevertheless suggests that working memory could be an important factor contributing to the ability to learn and use abstract categories.

Role of Language in Conceptual Development

Although there is little disagreement that language plays a critical role in conceptual development, what exactly this role is, how it changes in the course of development, and how it differs for different kinds of concepts are matters of debate. In particular, sometimes words denote already existing perceptual categories that are likely to be acquired in infancy (e.g., DOG, BALL, or CUP). Sometimes words are a starting point for forming nonperceptual categories (e.g., LOVE, FAIRNESS, or MEMORY). And sometimes language interacts with other aspects of experience to help form important ontological distinctions that are necessary for the development of conceptual hierarchies.

Learning Words for As-Yet-Unknown Categories. As discussed, there are many situations in which categories have enough statistical structure to enable them to be learned perceptually. Typically, these are statistically dense categories of objects, many of which are present in the environment surrounding the infant. In these situations, words are likely to follow category learning and thus are mapped onto these preexisting categories (Merriman, Schuster, & Hager, 1991; Mervis, 1987). What do words do if a child acquires a lexical entry for an already-known category (e.g., a word "dog" for a perceptual category DOG)? One possibility is that, at least initially, in these situations words function as features, thus simply contributing to the featural overlap among category members. Although I am not aware of any direct evidence supporting this contention, there is a growing body of indirect evidence.

First, there is evidence that shared labels contribute to similarity of the items (Sloutsky & Fisher, 2004a; Sloutsky & Lo, 1999; Sloutsky, Lo, & Fisher, 2001). In a number of studies, 4- to 5-year-olds were presented with a target and two test items and asked which of the test items looked more like the target. In one condition, there were no labels; in another condition, labels were introduced, such that one test item shared the label with the target (e.g., both were called "a dax") and another had a different label. The results indicated that items that shared the label were perceived as looking more similar than the same items introduced without labels. There is also more recent evidence (Sloutsky & Fisher, 2012) indicating that young children were more likely to infer that two items have similar properties when the items were accompanied by phonologically similar labels than when the items were accompanied by different labels. These effects should not have been observed if young children construed linguistic labels as symbols rather than as a feature of items.

Second, Deng and Sloutsky (2012) provided evidence that salient visual features have greater effects on category learning than do words. These researchers adopted a paradigm introduced by Yamauchi and Markman (1998, 2000) to distinguish between whether labels function as features or category markers. The paradigm is based on the following idea. Imagine two categories, A (labeled "A") and B (labeled "B"), each having five binary dimensions (e.g., Size: large versus small; Color: black versus white; etc.). Because the dimensions are binary, one value on each dimension can be denoted by "0" and another by "1" (e.g., white = 0, black = 1). Further, imagine the prototype of Category A has the value of "1" for all dimensions (i.e., "A," 1, 1, 1, 1, 1), and the prototype of Category B has the value of "0" for all dimensions (i.e., "B," 0, 0, 0, 0, 0). Items derived from these prototypes can be used in two interrelated generalization tasks—classification and projective induction. The goal of classification is to infer category membership (and hence the label) on the basis of presented features. For example, participants are first presented with all the values for an item, such that all the values except one come from Category A. Participants are then asked to predict the label (e.g., ?, 0, 1, 1, 1, 1). In contrast, the goal of induction is to infer a feature on the basis of category label and other presented features. For example, participants are given an item A with features 1, ?, 1, 0, 1 and asked to predict the value of the missing feature. A critical manipulation that could illuminate the role of labels is the "low-match" condition. For low-match induction, participants were presented with an item A as ?, 0, 1, 0, 0 (thus more similar to the prototype of Category B) and asked to infer the missing feature. For low-match classification, participants were presented with an item ?, 1, 0, 1, 0, 0 (which again was more similar to the prototype of Category B) and asked to infer the missing category label.

In both cases, items are more similar to prototype B, and, if labels are category markers, participants should be more likely to infer the missing feature as belonging to A (i.e., the induction task) than to infer label “A” (i.e., the classification task). In contrast, if the label is just another feature, then a different pattern should emerge: Relative performance on classification and induction tasks should depend on attentional weights of labels compared to those of other features. Specifically, if there are features with a higher attentional weight than the label, then a classification task (when a highly salient feature could be used to predict the label) should yield more A responses than an induction task (when the label is used to predict the highly salient feature). Deng and Sloutsky (2012) found that when all features were of comparable salience, 4- to 5-year-olds (in contrast to adults) tended to rely on the overall similarity rather than on category label. Furthermore, when the label was pitted against a highly salient visual feature (i.e., pattern of motion), 4- to 5-year-olds relied on the single most salient feature. Therefore, labels may function as features early in development, and they become category markers as a result of development.

Learning Words for Already-Known Categories. Even if words are features early in development, they do not have to remain features throughout development. First, there is evidence from the studies just cited indicating that adults are more likely to treat words as symbols rather than as features. Second, many concepts are learned in the order opposite to the one just described. That is, in contrast to the order of acquisition described (i.e., from prelinguistic categories to words), many concepts start with words. For example, around 4 years of age, a child may know words such as “love,” “number,” or “history” (MRC Psycholinguistic Database, http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm), but it is quite unlikely that the child knows the underlying concepts. Although I am unaware of any research examining this issue, it is hard to see how words can be features in these situations. It is more likely that, in these circumstances, words denote a category that is yet to be acquired.

Interaction Between Language and Other Aspects of the Experience. There is evidence that, as early as at 24 months, children exhibit an understanding of broad ontological distinctions, such as the distinction between objects and substances (e.g., Soja, 1992; Soja, Carey, & Spelke, 1991). How do children develop such understanding? Some have suggested that language (in the form of count/mass noun syntax) is instrumental in the acquisition of the ontological categories of object and substance (Quine, 1960); others have proposed that these broad ontological distinctions precede language and are thus independent of it (Soja et al., 1991; Soja, 1992). In contrast to these single-cause accounts, Smith and colleagues (e.g., Samuelson & Smith, 1999) proposed that perceptual cues (e.g., solidity) and linguistic cues (e.g., mass versus count noun syntax) jointly contribute to the acquisition of broad ontological distinctions. To test these ideas, they asked two interrelated questions. They asked whether solidity is correlated with syntax: Are solid things more likely to me labeled with count nouns and are substances more likely to be labeled by mass nouns? They also asked whether solidity is correlated with category organization: Are solid things more likely to be organized by shape and are nonsolid things more likely to be organized by material? To answer these questions, they selected a corpus of 312 nouns taken from the toddler form of the MacArthur Communicative Development Inventory (Fenson et al., 1994). They then asked adult participants to describe the solidity versus nonsolidity of items named by each noun and to describe the similarities in shape, material, and color of the instances named by each noun. Their findings are graphically presented in Figure 2.5. These results indicate that although syntax, solidity, and category structure do not overlap completely, there is a high level of correspondence among the three: Solids, unlike nonsolids, are more likely to be referred to by count nouns and to be organized by shape.

Development of Semantic Knowledge and Its Role in Conceptual Development

Although language is not a necessary aspect of category learning (nonhuman animals and prelinguistic infants can learn categories), lexicalization of categories is a critical step in acquiring and integrating knowledge about the world. First, language allows one to efficiently encode, store, and retrieve information about the category. Second, language allows one to acquire information that goes beyond one's own experience (e.g., owls are awake at night) or not observable directly (e.g., vegetables have vitamins). And third, language allows the establishment and communication of nontrivial commonalities (e.g., plants and animals are alike in that they need water to survive). Therefore, language allows the development of a conceptual network (also referred to as sematic knowledge) that represents one's knowledge about the world. Semantic memory is the system that stores semantic knowledge—information about concepts, facts related to these concepts, and words denoting them (cf. Tulving, 1972). Various tasks can be used to examine semantic memory, including picture naming, word-to-picture matching, sorting, category verification (e.g., Is a cat an animal?), and property verification (e.g., Do cats have wings?). The idea of semantic memory raises questions, such as these: How is knowledge represented in semantic memory? And how do these representations change in the course of development? Several proposals have been advanced to answer these questions.

One idea is that concepts are stored as nodes in a hierarchically organized network (Collins & Quillian, 1969; Quillian, 1967). Each node is linked to facts (or propositions) that are true of all (and only of) members of a given category and its constituent subcategories. Therefore, for example, facts stored about *canary* should be specific to canaries (but not necessarily to all birds) whereas facts stored about *the bird* should be specific to all birds.

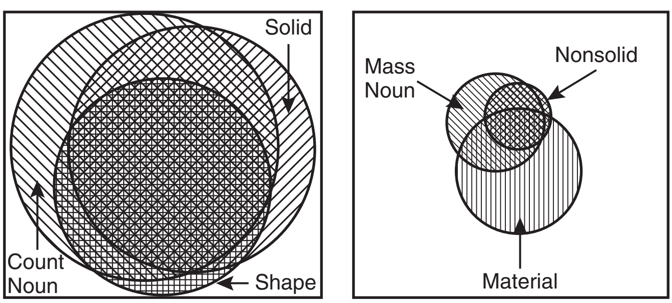


Figure 2.5 Codependency of solidity, shape, and syntax in early vocabulary.
SOURCE: After Samuelson & Smith (1999).

To test these ideas, Collins and Quillian (1969) presented adult participants with property verification sentences (e.g., "robins can fly") and category verification sentences (e.g., "a robin is a bird"). The authors predicted that people have faster access to information stored in a given node than to information stored in a superordinate node. As a result, participants should respond faster to category verification questions, such as "Is a canary a bird?" than to "Is a canary an animal?" They should also respond faster to property verification questions related to a particular level (e.g., "Can a canary sing?") than to those related to a superordinate level (e.g., "Does a canary have skin?"). All these predictions were confirmed empirically, thus suggesting that this model captures important properties of the human conceptual system.

Despite the early success of the model, subsequent researchers presented evidence that was difficult to reconcile with the model's predictions. (See Rogers & McClelland, 2004, for an extensive review.) First, contrary to the model predictions, reaction times in property-verification tasks were influenced by factors that had little to do with the position of the property in the taxonomic hierarchy (e.g., feature typicality and frequency). In addition, for many categories, the time it took to verify category membership differed from the model's predictions. Although closer higher-level categories should be identified faster than remote ones (e.g., the judgment that "X is a bird" should be faster than "X is an animal"), people are in fact faster to judge that a chicken is an animal than that it is a bird (Rips, Shoben, & Smith, 1973). And finally, multiple researchers (e.g., McClelland & Rogers, 2003; Rogers & McClelland, 2004; Rogers et al., 2004) noted that the model is inconsistent with neuropsychological literature on semantic disintegration (e.g., Warrington, 1975) as well as with literature on semantic development.

The second idea is that abstract semantic representations emerge as a product of domain-general statistical learning: Modality-specific perceptual representations provide the input to semantics, and modality-specific response systems permit the expression of semantic knowledge. For example, Rogers and McClelland (2004) examined semantic development using a variant of a connectionist network developed by Rumelhart and Todd (1993). The network learns propositions about the concepts. Input consists of a concept–relation pair (e.g., the input "Rose HAS"), and the network is trained to turn on all those output units that represent correct completions of the input pattern. Although the details of learning in the model are outside of the scope of this review, it is important to note that the network itself is feed-forward in that activation propagates forward but the error propagates backward using a variant of supervised learning known as the back propagation algorithm (Rumelhart, Hinton, & Williams 1986). In many variants of supervised learning, the system responds to a query and then receives feedback as to whether the response is correct or not. Learning is construed as the process of error reduction, and back propagation is a formal way of reducing the error.

A critical component of the model is the idea of coherent covariation, that is, co-occurrence of a set of properties across different category members. Coherent covariation is distinct from simple correlation in that it generally refers to the co-occurrence of multiple rather than just two properties. For example, having wings, having feathers, having beaks, living in nests, and having hollow bones all consistently co-occur in birds. The model accounts for a variety of developmental data, most important, for progressive differentiation of concepts in the course of development. Progressive differentiation is the idea that broader categorical distinctions (e.g., the distinction between animates and artifacts) is acquired prior to more specific categorical distinctions (e.g., the distinction between cats and dogs).

The Rogers and McClelland model offers a mechanistic account of semantic development, makes clear theoretical predictions, and explains some of the best known developmental findings. However, many theoretical ideas advanced by Rogers and McClelland are yet to be tested in empirical studies. In particular, it will be important to systematically measure developmental changes in the structure of semantic memory and to examine whether the model captures these changes. It is also worth mentioning that progressive differentiation is the only mechanism of semantic development captured by the model. It is not clear whether this mechanism is capable of learning abstract concepts (e.g., legal, scientific, or mathematical) that combine items that have few commonalities (and thus require the learner to ignore differences between instances of a concept).

The third idea is that a semantic network can be described as a graph that consists of a set of nodes and a set of edges that connect individual nodes (e.g., Hills, Maouene, Maouene, Sheya, & Smith, 2009; Steyvers & Tenenbaum, 2005). An extensive treatment of the graph-theoretical approach is presented elsewhere (Steyvers & Tenenbaum, 2005), and I consider here only the basic concepts of the approach and the ways it captures development. According to this idea, two connected nodes are considered to be neighbors, and a node and all its neighbors are considered a neighborhood. The approach allows for a number of quantitative measures, including the size of the network (i.e., the number of nodes) and the clustering coefficient. The latter is determined by calculating the number of connections between the nearest neighbors of a given node and the total number of possible connections. Hills et al. (2009) used this approach to examine the network of nouns that were learned early by 2.5-year-olds.

The goal of the network analysis was to answer two questions: how well toddlers' basic-level concepts are organized into superordinate categories and how perceptual and conceptual (i.e., functional) features contribute to that organization. To answer the first question, the resulting networks were analyzed by calculating the clustering coefficient, which was then compared to a control random network with an equivalent number of nodes. It was found that when concepts shared few features, there was only one densely connected network: Everything was connected to everything else, and no structure emerged. In contrast, when concepts shared multiple features, structure was more apparent: The resulting network had clusters of nouns representing animals, vehicles, foods, clothes, and household objects. Therefore, structure may emerge in the course of development as children learn multiple properties shared by related nouns. This approach also offers an interesting possibility for studying the development of knowledge domains. As conceptual neighborhoods become increasingly more coherent and increasingly distinct from other neighborhoods, they may evolve into what is known as knowledge domains.

Network analyses conducted to answer the second question indicated that perceptual features are more redundant and provide robust information about category inclusion whereas conceptual features are rarer and provide a better discrimination between categories. "A single conceptual relation is sufficient to define all category members that are, for example, *used for transportation*. No single perceptual feature contains that information" (Hills et al., 2009, p. 389). Therefore, both perceptual and conceptual features were found play important and perhaps complementary roles in early conceptual organization.

Origins of Semantic Knowledge

Although the accounts just reviewed suggest that semantic knowledge emerges from the learner's interactions with the world, no comprehensive account has yet been offered about how the abstract predicates (e.g., "ISA," which reflects a relation of class inclusion in Rogers & McClelland, 2004, or functional features, such as "used for transportation" in Hills et al., 2009) emerge from nonconceptual primitives.

Although a complete account is lacking, there are a number of partial accounts. As will be discussed, some have argued that semantic knowledge emerges from experience; others have argued that components of semantic knowledge exhibit early onset and are unlikely to stem from individual experiences. One attempt to explain the development of semantic relatedness by linking it to experience was offered by Fisher and colleagues (Fisher, 2010; Fisher, Matlen, & Godwin, 2011). These researchers examined the development of semantic relatedness by presenting participants with verbal inductive arguments. For example, upon being told that dogs have property X, will participants generalize this property to semantically related items, such as puppies? The investigators selected semantically related (SR) items that were highly familiar to even the youngest participants (verifying this familiarity in a separate experiment). In addition, they also established through the analysis of Child Language Data Exchange System (CHILDES) corpus that some of the SR items tended to co-occur in the same sentence (e.g., bunny-rabbit) while others were unlikely to co-occur (e.g., crocodile-alligator). The results indicated that 4-year-olds generalized properties only when the SR items were co-occurring (e.g., from bunny to rabbit but not from crocodile to alligator). In contrast, 5- and 6-year-olds generalized even when SR items were not co-occurring. Therefore, between 4 and 6 years of age, children undergo semantic development, and this development affects their pattern of inductive inference. So, what develops between 4 and 6 years of age? Perhaps children develop a more coherent taxonomy of their concepts and a better mapping of words on this taxonomy (cf. Nelson, 1974, for related arguments). Or perhaps some other changes are at the heart of semantic development. A detailed developmental account of these findings is yet to be provided.

There is also an argument that some components of semantic knowledge are unlikely to stem from individual experience. For example, some argue that even young children attach special significance to information presented in the "generic" format (Cimpian & Erickson, 2012; Cimpian & Park, 2014). The generic format (e.g., *Dogs bark*) involves a statement that has an omitted existential quantifier "some" and thus should be equivalent to the statement *Some dogs bark*. However, research indicates that this format may be doing something different from existential quantification. In particular, it has been argued that even young children place special value on generic information, often inferring that it provides important insights about the world. For example, Cimpian and Scott (2012) presented 4- to 7-year-olds novel facts that were in either generic format (e.g., *Hedgehogs eat hexapods*) or nongeneric format (e.g., *This hedgehog eats hexapods*). Children were then asked whether other people (e.g., their parents or grown-ups in general) knew these facts. It was found that children were more likely to expect adults to know facts that had been presented in the generic format. Cimpian and Markman (2009) also reported that features presented in the generic format were more likely to be construed as causal. Although the mechanisms of the effect of generic format is not known, it is possible that people (including young children) interpret it as a universally quantified statement (e.g., *All X are Y*), suggesting that the statement describes the entire class. (See Cimpian & Erikson, 2012.) However, the effect is so far construed as reflecting “a generic bias,” which appears to be closer to describing rather than explaining the effect.

Development of Conceptual Hierarchies

One hallmark of conceptual organization is that it has a structure, and taxonomic organization of categories is an example of such structure. Although taxonomies are not the only possible structure (see Kemp, Shafto, & Tenenbaum, 2012, for discussion of other possibilities), it is perhaps the most general and well-studied one. An example of such taxonomic hierarchy is Fido → Dog → Mammal → Animal → Living Thing → Bounded Thing → Thing. It is clear that such hierarchies are based on class-inclusion relations—they require including a set of mutually exclusive lower-level categories A_i into a higher-level category B. For a system to be a hierarchy, it has to satisfy two important constraints. First, lower-level categories should be exhaustive with respect to a higher-level category, such that $A_1 + A_2 + \ldots + A_i = B$. In practice, if not all subcategories are known, the exhaustiveness can be achieved by dividing B into A and its complement A’, such that A + A’ = B (e.g., animals consist of cats and non-cat animals). The second constraint is that subclasses of B should be mutually exclusive, that is, they should have no common members (i.e., the intersection of the two sets should be equal to 0: A ^ A’ = ∅). It seems that a number of abilities should be in place in order for a taxonomic organization of concepts to be possible. First, there should be an appreciation of the logical constraints (e.g., understanding of the fact that the subclasses have to be mutually exclusive and that they are properly included in a larger class). Understanding of class inclusion relations manifests itself in understanding of quantifiers, such as *All*, *Some*, *Some are not*, and *None*. Second, there should be knowledge of words denoting higher classes: Although a lower-level class can be derived from a higher-level class by using an adjective (*dog* + *adj* (small) = *small dog*), a higher-order class for a dog cannot be derived and requires knowledge. And third (somewhat related to the second point), there should be knowledge of a domain in which a taxonomy is to be built. In the absence of such knowledge, it may not be clear which entities form categories and which categories are bound with class-inclusion relations and which are not. Of course, these abilities do not have to emerge all at the same time. Therefore, each of these abilities may represent a starting point for the development of conceptual hierarchies. Historically, a variety of candidate starting points have been considered. Some have argued that the development of conceptual hierarchies starts with logic, some argued that it starts with language, and some argued that it starts with domain knowledge.

Logic of Classes as a Starting Point

In their classic book on the development of classification, Inhelder and Piaget (1964) considered the development of conceptual hierarchies as a function of the development of the logic of classes. The idea of the logic of classes is that multidimensional sets of stimuli can be divided into proper subsets focusing on one dimension at a time, especially when dimensions are fully crossed. Therefore, as shown in Figure 2.6, set S can be divided according to dimension 1 into two

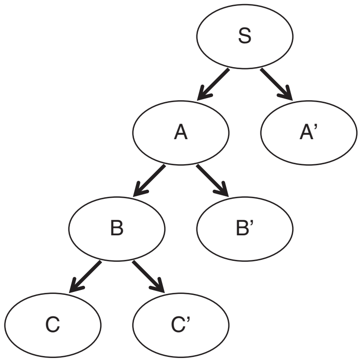


Figure 2.6 Simple hierarchy-based class-inclusion relations.

mutually exclusive classes (e.g., Red objects *A* and Non-Red objects *A'*). A can further divided into subsets *B* (e.g., angular objects) and *B'* (nonangular objects), which in turn can be further divided in *C* (squares) and *C'* (nonsquared angular objects). Fundamental changes occur with respect to understanding of class inclusion relations, and once these relations are mastered, a classification scheme based on these relations can be applied to any domain of knowledge. However, it easy to notice that logic alone may not be sufficient for building such hierarchies. In addition to logic, one needs to know dimensions that distinguish subcategories, which may be a nontrivial task. For example, a division of objects into black and white ones is trivial, but a division of animals into feline and canine animals may be not as trivial. Therefore, most contemporary theories consider domain knowledge as a necessary component of the development of conceptual hierarchies.

Domain Knowledge Approach

As noted by Chi, Hutchinson, and Robin (1989), "[I]n many instances, having knowledge in a specific domain can overcome any limitations that could have been imposed by the lack of global operators. Yet lacking knowledge in a specific domain also can prevent adults from reasoning logically, even though they are presumed to have the logical operators" (p. 28). Obviously, the same logical structure (i.e., class inclusion) may be based on different properties, and these properties may differ in their support of category coherence and inductive inference. For example, the property "has small parts" provides much weaker support for inductive inference than the property "has gills." In addition, lower-level categories may share few attributes with higher-level categories, or they may share many attributes. (The same is true for individuals with respect to categories.) The latter structure will result in greater coherence than will the former. A number of researchers (Carey, 1985; Chi et al., 1989; Inagaki & Hatano, 2002; Keil, 1981) subscribe to the view that a hierarchical organization of concepts may result from knowledge of a domain. In this case, class inclusion relations simply follow from a structural representation of a domain, without necessarily reflecting a more general ability to honor class inclusion. For example, mere knowledge of dinosaurs may help the child understand that all brontosauri are dinosaurs, but not all dinosaurs are brontosauri, without necessarily enabling the child to apply class-inclusion relations to unknown domains.

More recently, the assumption of hierarchical knowledge preceding the development of logic was used in a model of word learning proposed by Xu and Tenenbaum (2007). The model construes word learning as a variant of Bayesian inference and attempts to explain how young word learners select a referent for a newly learned word. For example, if a child is shown a terrier and it is called "a dax," the child, according to Xu and Tennenbaum, needs to decide whether the word refers to terriers, dogs, or all animals. However, in contrast to other domain knowledge approaches, this model presumes a very early emergence of a conceptual hierarchy, which raises the question of where this hierarchy itself came from.

Role of Language and Parental Input in the Development of Taxonomic Hierarchies

Is it possible that language cues help children form hierarchies? A number of studies addressing this issue have provided limited support for this idea. For example, Callanan (1985, 1989) examined whether the ways categories are labeled may affect children's interpretation of the referent class. It turned out that, when introducing new words referring to the superordinate level, parents are likely to anchor these at the basic level. In particular, when introducing the word "animal" (i.e., a superordinate category), a parent may point to a dog (i.e., a basic-level category) and say: "Here is a dog; it is a kind of animal." However, despite these strategies, 3- to 4-year-old children are highly unlikely to interpret new words as referring to superordinate categories (Callanan, 1989). Overall, evidence suggests that, at least for preschoolers, (a) spontaneous categorization at the superordinate level is rather rare and (b) parents rarely name items at the superordinate level.

Unresolved Issues

Although it is likely that people eventually form conceptual hierarchies, the process of development is not well understood. Some (e.g., Piaget; see Inhelder & Piaget, 1964) argued for protracted (yet spontaneous) development, which is not fully completed until the stage of concrete operations or perhaps even later. Others argued that this ability transpires significantly earlier, with many preschoolers exhibiting evidence of conceptual hierarchies. However, evidence for the early onset of conceptual hierarchies is limited. Most important, even if a child exhibits the ability to classify items at a superordinate level or draws inductive inferences on the basis of a superordinate class, this ability does not necessarily indicate the presence of a conceptual hierarchy (cf. Halford et al., 2002). This is because these classifications or inferences may be driven by similarity (i.e., members of the same superordinate category are more similar to each other than to nonmembers) rather than by their place in a conceptual hierarchy. It seems that a critical prerequisite of a conceptual hierarchy is the understanding of class inclusion, and this understanding may be missing early in development (e.g., Greene, 1994; Siegler & Svetina, 2006; Winer, 1980).

The second issue concerns factors affecting the development of conceptual hierarchies. Although there is a widely shared expectation that the development of conceptual hierarchies is spontaneous, there is little evidence that (at least early in development) parents label items at the superordinate level or attract children's attention to superordinate classes (Blewitt, 1983; Callanan, 1985, 1989). Therefore, it is possible that conceptual hierarchies are a consequence of formal education (Scribner & Cole, 1973). Although both issues remain unresolved, it seems fairly clear that the development of conceptual hierarchies in a given domain is based on at least two prerequisites: (1) understanding of class-inclusion relations and the logic of quantification and (2) knowledge of how these relations can be applied in a particular domain.

Role of Categories in Inductive Inference

There is general agreement that one of the central functions of categories is to subserve prediction (e.g., Anderson, 1991). Therefore, it is hardly surprising that the ability to draw inductive inferences has been used to probe conceptual development. Although several researchers have presented evidence for the ability of infants to perform induction, the majority of research on inductive inference focuses on verbal children. Several questions appear to be critical: What is the mechanism of early induction and how does it change in the course of development? To what extent does prior knowledge constrain inference? How flexible is the inference? And what is the role of words in inductive inference?

Mechanism of Early Induction

Although it is well established that induction appears early in development (Gelman & Markman, 1986; Mandler & McDonough, 1996; Sloutsky & Fisher, 2004a; Welder & Graham, 2001), the mechanism of early induction remains unclear. In an attempt to understand early induction, two theoretical proposals have been formulated: the knowledge-based approach and the similarity-based approach.

According to the first approach, early induction is a two-step process: First, people (including young children) identify the category of an entity and then generalize properties of the entity to other members of the category. Therefore, if told that a dog has a certain biological property (e.g., a particular type of heart) and then asked to generalize this property (e.g., "Who is more likely to have the same heart, another dog or a cat?"), people generalize the property to another dog because the two dogs belong to the same category. Therefore, even early in development, induction is said to be *category-based*. The ability to perform category-based induction hinges on a number of assumptions attributed to young children. Most important, young children are expected to hold the *category assumption*—a belief that individuals belong to general categories, with members of the same natural kind category sharing many important properties. In addition, young children are expected to hold the *linguistic assumption*—a belief that count nouns denote categories. Although it is not claimed that these assumptions are part of children's explicit knowledge, it is generally argued that early induction is based on them.

Support for the idea that early induction is category-based comes from several sources. First, in a series of experiments, Gelman and Markman (1986) presented young children with a triad task, in which stimuli consisted of one target and two test items. The triad task was designed to pit appearance similarity against category membership: One test item belonged to the same category as the target but looked dissimilar from the target; the other test item looked similar to the target while belonging to a different category. Participants were presented with a triad and were informed that one test item had a particular hidden property (e.g., "hollow bones") while the other test item had a different hidden property (e.g., "solid bones"). The task was to generalize a hidden property to the target. Category membership was communicated by using the same label for the target and the dissimilar test item. In general, children were more likely to generalize the property of the test item that shared the target's label than the property of the test item that shared the target's appearance. (But see Sloutsky & Fisher, 2004a, Experiment 4, for diverging evidence and counterarguments.) This finding was interpreted as evidence that children's induction is based on common category information.

According to the *similarity-based* approach, induction starts out as similarity-based and becomes category-based as a result of development. Although it is not known precisely when induction becomes category-based, proponents of this approach argue that early induction is the same process as early categorization, with both being based on computing similarity between a presented item (or an item stored in memory) and a to-be-judged item.

Although proponents of both positions expect linguistic labels to affect induction, the mechanisms assumed to drive these effects differ radically between these positions. According to the knowledge-based approach, labels affect induction because they denote category membership, with category information driving induction. According to the similarity-based approach, labels affect induction because they contribute to the perceived similarity of items, with similarity driving induction. Therefore, evidence that children rely on a category label in a triad induction task is not sufficient for distinguishing between the two positions.

One way of deciding whether induction is category-based or similarity-based is to examine memory traces formed during an induction task (Sloutsky & Fisher, 2004a, 2004b; see also Hayes & Heit, 2004, for a review). The idea is based on the following reasoning. There is a well-known "level-of-processing effect" in which deeper semantic processing facilitates memory so that there is better recognition of presented items (i.e., a higher proportion of "hits"; see Craik & Lockhart, 1972; Craik & Tulving, 1975). Several recent studies, however, indicate that deeper processing results not only in higher hit rates but also in more memory intrusions (i.e., false recognitions of nonpresented items that are "critical lures," or items that are semantically associated to the original items; e.g., Rhodes & Anastasi, 2000; Thapar & McDermott, 2001). It also has been demonstrated that when to-be-remembered items are related categorically, participants often produce false alarms by falsely recognizing critical lures that are nonpresented members of studied categories (Koutstaal & Schacter, 1997). And it is known that focusing on perceptual details of pictorially presented information leads to more accurate recognition (Marks, 1991). Although hits in this case might be slightly lower, false alarms are significantly lower than when participants are engaged in deep semantic processing. Collectively, these findings suggest that categorization (which is a variant of deeper semantic processing) would result in a higher level of memory intrusions and thus in lower recognition accuracy than shallow perceptual processing. (See also Brainerd, Reyna, & Forrest, 2002, for related arguments.)

Thus, a memory test administered after an induction task may reveal differential encoding of information during induction: if participants perform category-based induction, they should be engaged in deep semantic processing and therefore exhibit low discrimination of studied items from critical lures during a memory test (compared to a no-induction baseline condition). If, however, participants perform similarity-based induction, they should be engaged in shallow perceptual processing, and, as a result, their memory accuracy should not decrease compared to the baseline. Because young children, unlike adults, were expected to perform similarity-based induction, this reasoning led to a nontrivial prediction that after performing induction, young children may exhibit greater memory accuracy (i.e., have fewer false alarms) than adults.

These predictions have received empirical support: The pattern of results reported by Sloutsky and Fisher (2004a, 2004b) indicates that while adults perform category-based induction, young children perform similarity-based induction. In particular, after performing inductive generalizations about members of familiar animal categories (i.e., cats, bears, and birds), adults' memory accuracy attenuated markedly compared to the no-induction baseline. At the same time, young children were accurate in both the baseline and induction conditions, exhibiting greater accuracy in the induction condition than adults. However, after providing short training on category-based induction (participants were taught that things that have the same name belong to the same kind and have much in common), memory accuracy of 5-year-olds decreased to the level of adults in the induction condition. At the same time, training did not attenuate children's accuracy in the baseline condition. That is, even after training, 5-year-olds exhibited high accuracy on recognition memory tasks. These findings suggest that the decrease in memory accuracy observed in the induction condition is attributable to the specific effects of training to perform category-based induction rather than to general factors such as fatigue. These results demonstrate that young children (unlike adults) spontaneously perform induction in a similarity-based rather than category-based manner and that they can learn to perform category-based induction via simple training. In a subsequent study, Fisher and Sloutsky (2005) demonstrated that category-based induction undergoes protracted development, with recognition memory accuracy dropping to the level of adults only by 11 years of age (see Figure 2.7). The development of category-based induction is inferred from the semantic interference effect, that is, from lower memory in the induction condition than in the baseline condition.

Another way of examining the mechanism of inductive inference was suggested by Sloutsky, Koos, and Fisher (2007), who gave participants direct access to category information by teaching them a new natural-kind category that had a clear category-identification rule. Once participants had learned the category, they were presented with an induction task, in which category membership was pitted against appearance. If, for natural kinds, category-based induction is the default, then young children (who successfully learn the category) should assume that members of the same kind have much in common. As a result, when performing induction, they should rely on category membership and ignore appearance information. Conversely, if similarity-based induction is the default, then young children (even when they successfully learn the category) should rely on appearance

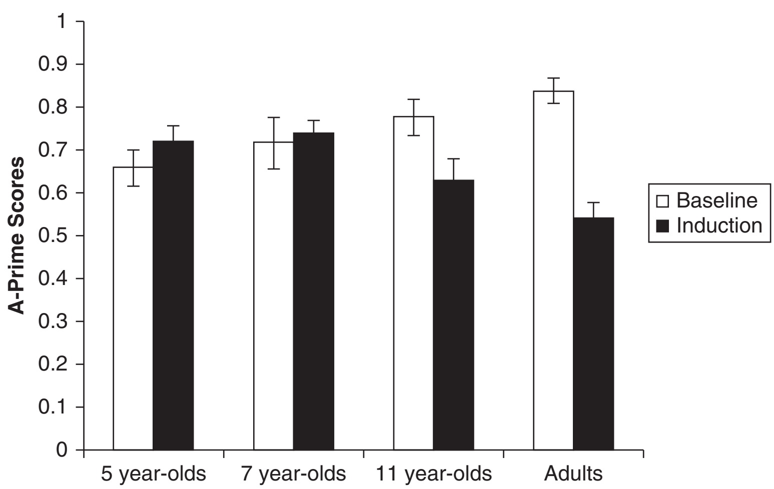


Figure 2.7 Development of category-based induction.
SOURCE: After Fisher & Sloutsky (2005).

information while disregarding category membership information.

In the experiments reported by Sloutsky et al. (2007), 4- to 5-year-olds were first presented with a category learning task during which they learned that artificial animal-like creatures belong to two natural kinds: nice, friendly pets or wild, dangerous animals. The membership in a category could be detected by a rule; appearances were not predictive of category membership. Children were then given a categorization task with items that differed from those used during training. Participants readily acquired these categories and accurately sorted the items according to their kind information. Then participants were presented with a triad induction task. Each triad consisted of a target and two test items, with one test item sharing the target's category membership but not its appearance and the other test item sharing the target's appearance but not its category membership. Participants were familiarized with a quasi-biological property of the target and asked to generalize this property to one of the test items. Finally, participants were given a final (i.e., postinduction) categorization task using the same items as in the induction task. The results provided little support for category-based induction early in development: 4- to 5-year-olds successfully learned the categories but generalized properties on the basis of common appearance. (See Figure 2.8.)

One potential criticism of this research is that the researchers failed to communicate conceptual information to young children. As a result, children might have interpreted these categories as artificial groupings rather than natural kinds that support inductive inference. There are several reasons to believe that this criticism is wrong. First, Sloutsky et al. (2007) communicated the biological relevance of the category-defining information and consistently referred to the studied categories as "kinds of animals." More important, unpublished data by Sloutsky et al. and recent published data by Gelman and Davidson (2013) suggest that adults interpreted these categories as natural kinds and based their induction on these categories. Therefore, nothing in the description suggests that the categories are not natural kinds. However, it is also possible that although information provided by the researchers was sufficient for adults to infer that the studied categories were natural kinds, it was not sufficient for young children.

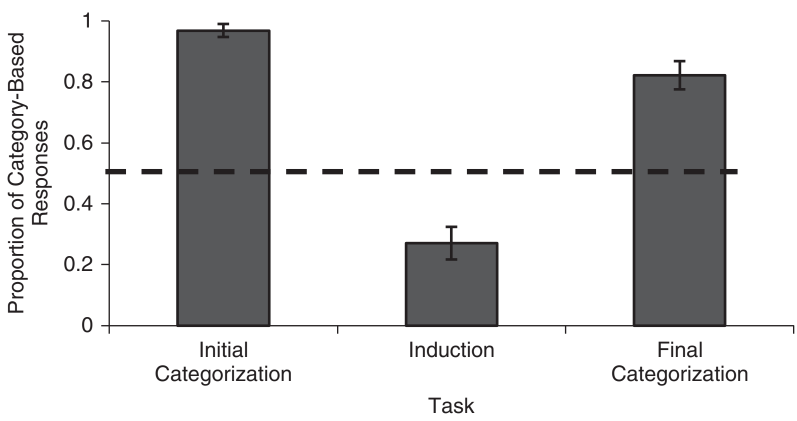


Figure 2.8 Proportions of category-based responses by task in Sloutsky et al. (2007). The dashed line indicates chance responding.

Recently, Gelman and Davidson addressed this possibility by making every effort to communicate to young children that the categories were indeed natural kinds. They found that under these conditions, 4- to 5-year-olds did perform category-based induction with the newly learned categories. However, as argued by Sloutsky et al. (2015), Gelman and Davidson (2013) changed many other aspects of the original study as well (e.g., they made the category-defining information highly salient and used a training regime that could have attracted attention to this highly salient information). Sloutsky et al. (2015) went on to demonstrate that these attentional manipulations rather than conceptual information directed children's attention to category-defining information such that they subsequently used this information in their induction. Once these attentional factors were eliminated, effects of conceptual information were negligible. However, although 4- to 5-year-olds did not rely on conceptual information, 6- to 7-year-olds were more likely to do so, although significantly below the levels of adults. Therefore, category-based induction seems to be a product of development, and understanding this development in greater detail is a task of future research.

Development of Inductive Inference

Many models of inductive inference view generalization as the result of computing the overlap or similarity between the features of the premise (or inductive base) and the conclusion (e.g., Osherson, Smith, Wilkie, López, & Shafir, 1990; Sloman, 1993; Sloutsky & Fisher, 2004a). Therefore, whether the items are presented as pictures or as verbal arguments, people are generally more likely to generalize a property from a robin to a blue jay than from a robin to a monkey. Although most researchers agree that premise-conclusion similarity is important, some argue that category information is important as well. For example, Osherson et al. (1990) in their influential Similarity-Coverage model of induction focused on two components that potentially guide induction: the similarity component (which reflected the premise-conclusion similarity) and the coverage component. The coverage component focuses on how well the premise category covers the conclusion category. For example, in the argument "Mice and bears have an ulnary artery, therefore mammals have an ulnary artery," premise categories (i.e., mice and bears) provide broad coverage of the conclusion category (i.e., mammal). In contrast, in the argument "Mice and rats have an ulnary artery, therefore mammals have an ulnary artery," premise categories provide narrow coverage of the conclusion category.

Several phenomena are diagnostic of the coverage component, with *monotonicity* and *diversity* being most extensively studied in developmental literature. Monotonicity reflects the effect of sample size on induction. For example, the inference from *robins, eagles, and sparrows* to *birds* is stronger than the inference from *robins* to *birds*. Diversity reflects the effect of sample variability on induction. For example, the inference from *robins, falcons, and chicken* to *birds* is stronger than the inference from *eagles, hawks, and falcons* to *birds*. The coverage component seems to reflect the extent to which induction is category-based. What is the developmental time course of category-based induction as reflected in the development of the coverage component?

A number of studies (e.g., Gutheil & Gelman, 1997; López. Gelman, Gutheil, & Smith, 1992; Rhodes, Gelman, and Brickman, 2010) focused on monotonicity and diversity in an attempt to examine the development of category-based induction. The results indicate that although adults make use of information concerning sample size (larger samples are a stronger basis of inference than are smaller samples) and sample diversity (more diverse samples are better than more homogeneous samples) when making category-based inductive judgments, children do not do so until age 8 or 9 and even then to only a limited degree. These results converge with findings (e.g., Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004a, 2004b) suggesting a protracted development of category-based induction.

However, a number of more recent studies suggest that the development of the coverage component may occur earlier than previously believed. In one study, Rhodes et al. (2010) compared sensitivity to sample diversity in 5-year-olds and adults under two conditions. In the Expert condition, properties of the premise animals were communicated by an expert (a character who was introduced as knowing a lot about animals); in the Novice condition, these properties were communicated by a novice character who was introduced as having recently discovered these properties. In addition, in contrast to the previous research, both premise and conclusion categories were instantiated with pictures. Therefore, a nondiverse premise set included pictures of three Dalmatians and the conclusion was a picture of a collie. In contrast, a diverse premise set included a Dalmatian, a golden retriever, and a basset hound, and the conclusion was again the collie. Surprisingly, in the expert condition, 5-year-olds were very similar to adults in that they were much more likely to generalize on the basis of a diverse sample. However, in the novice condition, 5-year-olds exhibited an unexpected pattern (see Figure 2.9): Although 5-year-olds' reliance on diverse arguments did not decrease, their reliance on nondiverse arguments increased dramatically. These are provocative findings, and they raise several questions. First, given relatively strong reliance on diverse premises in both expertise conditions and given that

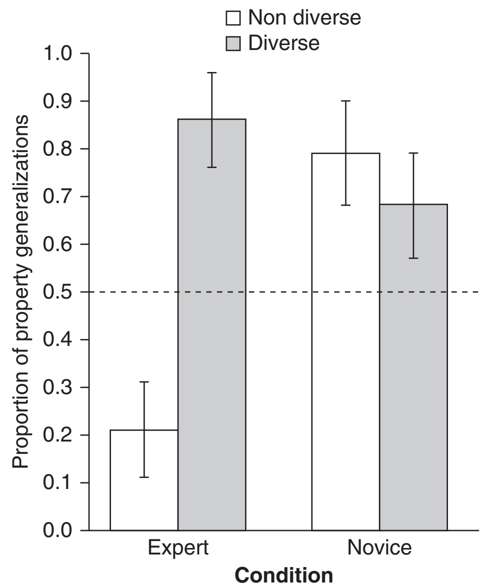


Figure 2.9 Proportions of property generalizations to basic-level matches, by condition and sample in 5-year-olds. The dashed line indicates chance responding.
Source: After Rhodes et al. (2010).

the premises were instantiated with pictures, is it possible that many premise pictures merely increase premise conclusion similarity compared to the previous studies? And second, why did the novice condition result in increased reliance on nondiverse premises?

In another study, Hayes and Thompson (2007) examined the development of sensitivity to potentially causal relations between a premise feature and conclusion feature (e.g., has large eyes → can see in the dark). Obviously, reliance on a causal connection between the premise and conclusion category is a more advanced form of inductive inference than reliance on similarity. Children (aged 5, 8, and 9 years) and college undergraduates were presented with two new categories, Waddo and Xoxney, and a description of each category. The description included three features including two that had a potential causal connection (e.g., "has large eyes" and "can see in the dark"); the third feature was unrelated to the other two (e.g., “has white wings”). Then participants were presented with an induction test in which the target was described as having a causal feature of Waddo (e.g., “has large eyes”) and a noncausal feature of Xoxney (e.g., “has a long beak”). Participants were then asked if the target could see in the dark like Waddo or jump high like Xoxney. It was reasoned that if participants understand the causal connection, they should systematically select Waddo; otherwise, their responding would be at chance. The result indicated that when causal information was made explicit, even 5-year-olds were above chance in relying on it. However, when it was not made explicit, even 8- to-9-year-olds were at chance. Therefore, it is not clear what drives the effects: Is it causal relatedness, or is it any link between or among features? Fortunately, the authors addressed this question in a separate experiment in which they first explicated causal relations (e.g., “they have *large eyes* to *better see in the dark*”) as well as noncausal temporal relations (e.g., “they *touch the bark* when they *eat leaves from trees*”). They then pitted a causal feature (“has large eyes”) and a noncausal feature (“touch the bark”) and asked to predict whether it sees in the dark like Waddo or eat leaves from trees like Xoxney. In this condition, 5-year-olds were at chance, but older children and adults tended to rely on causal features. This research suggests that 5-year-olds rely on any correlated features, and 8- to-9-year-olds rely on causally related features. Therefore, reliance on deeper properties and theoretically important relations in the course of induction is a result of protracted development. Although the factors contributing to these changes are not known, given how protracted the development is, it is likely that formal education is a contributing factor. However, this is merely a conjecture, and extensive research is needed to evaluate this hypothesis.

Summary

Concepts undergo dramatic development after infancy. First, there are developments that are likely to be attributed to more general cognitive development, including the development of attention and memory. In particular, children develop the ability to acquire increasingly sparse categories, thus becoming less dependent on similarity and within-category featural overlap. Second, language becomes an important source of conceptual development, with many concepts (e.g., LOVE, MATTER, or NUMBER) originating in language. Furthermore, acquisition of quantifiers may contribute to the development of mastery of class-inclusion relations. Third, there is evidence of a semantic development, with concepts forming conceptual networks of increasing within-network coherence and between-network differentiation. These networks may give rise to knowledge domains, reflecting the structure (taxonomical or otherwise) of these domains. And fourth, conceptual networks give rise to category-based inference, supplementing the earlier-emerging ability to perform inductive generalization on the basis of similarity. Each of these developments is likely to involve different processes and mechanisms, and the goal of future research is to uncover these processes and mechanisms.

CONCLUSION

This chapter provided an overview of conceptual development from infancy onward. The chapter formulated several principles of conceptual development, including (1) the diversity of conceptual behaviors; (2) the greater universality and earlier onset of simpler forms of conceptual behavior; (3) the development of more complex forms of conceptual behavior on the foundation of the simpler forms and the dependence of these more complex forms on other aspects of cognitive and language development; (4) the importance of the structure of input for learning; and (5) the developmental progression from less structured representations of concepts to more structured representations. The subsequent review attempted to present evidence for these principles.

The vast literature on conceptual development does not mean that our understanding of conceptual development is complete. As discussed in this chapter, many aspects pertaining to the origins of the ability to acquire categories, factors driving its development, and the underlying neurobiology are not well understood. Meeting the challenge of understanding developmental, cognitive, and brain mechanisms of conceptual behaviors will result in deeper, more complete understanding of the ability so central for our intelligence.

REFERENCES

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*, 409–429.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3–26.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–609.

Behl-Chadha, G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition*, *60*, 105–141.

Bhatt, R. S., & Quinn, P. C. (2010). How does learning impact development in infancy? The case of perceptual organization. *Infancy*, *16*, 2–38.

Blewitt, P. (1983). Dog versus collie: vocabulary in speech to young children. *Developmental Psychology*, *19*, 602–609.

Bomba, P. C., & Siqueland, E. R. (1983). The nature and structure of infant form categories. *Journal of Experimental Child Psychology*, *35*, 294–328.

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development*, *80*, 15–22.

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development*, *73*, 1363–1377.

Callanan, M. A. (1985). How parents label objects for young children: the role of input in the acquisition of category hierarchies. *Child Development*, *56*, 508–523.

Callanan, M. A. (1989). Development of object categories and inclusion relations: preschoolers' hypotheses about word meanings. *Developmental Psychology*, *25*, 207–216.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616.

Casasola, M., & Bhagwat, J. (2007). Do novel words facilitate 18-month-olds' spatial categorization? *Child Development*, *78*, 1818–1829.

Casasola, M., Bhagwat, J., & Burke, A. S. (2009). Learning to form a spatial category of tight-fit relations: how experience with a label can give a boost. *Developmental Psychology*, *45*, 711–723.

Cimpian, A., & Erickson, L. C. (2012). Remembering kinds: new evidence that categories are privileged in children's thinking. *Cognitive Psychology*, *64*, 161–185.

Cimpian, A., & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: evidence from their open-ended explanations. *Cognition*, *113*, 14–25.

Cimpian, A., & Park, J. J. (2014). Tell me about pangolins! Evidence that children are motivated to learn about kinds. *Journal of Experimental Psychology: General*, *143*, 46–55.

Cimpian, A., & Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition*, *123*, 419–433.

Chi, M. T. H., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly*, *35*, 27–62.

Cohen, L. B., & Cashon, K. H. (2006). Infant cognition. In W. Damon & R. Lerner (Series Eds.) & D. Kuhn & R. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed., pp. 214–251). Hoboken, NJ: Wiley.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 240–247.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: a role for associative learning. *Psychological Review*, *112*, 347.

Cowan, N. E. (1997). *The Development of Memory in Childhood*. Hove, UK: Psychology Press/Erlbaum.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*, 125–127.

Deng, W., & Sloutsky, V. M. (2012). Carrot eaters or moving heads: inductive inference is better supported by salient features than by category labels. *Psychological Science*, *23*, 178–186.

Deng, W. S., & Sloutsky, V. M. (2015a). The development of categorization: effects of classification and inference training on category representation. *Developmental Psychology*, *51*, 392–405.

Deng, W. S., & Sloutsky, V. M. (2015b). Linguistic labels, dynamic visual features, and attention in infant category learning. *Journal of Experimental Child Psychology*, *134*, 62–77.

Deng, W. S., & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62. http://doi.org/10.1016/j.cogpsych.2016.09.002

Eimas, P. D. (1994). Categorization in early infancy and the continuity of development. *Cognition*, *50*, 83–93.

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, *65*, 903–917.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*, 1–173.

Fisher, A. V. (2010). What's in the name? Or how rocks and stones are different from bunnies and rabbits. *Journal of Experimental Child Psychology*, *105*, 198–212.

Fisher, A. V., Godwin, K. E., Matlen, B. J., & Unger, L. (2015). Development of category-based induction and semantic knowledge. *Child Development*, *86*, 48–62.

Fisher, A. V., Matlen, B. J., & Godwin, K. E. (2011). Semantic similarity of labels and inductive generalization: taking a second look. *Cognition*, *118*, 432–438.

Fisher, A. V., & Sloutsky, V. M. (2005). When induction meets memory: evidence for gradual transition from similarity-based to category-based induction. *Child Development*, *76*, 583–597.

Fodor, J. (1972). Some reflections on L.S. Vygotsky's thought and language. *Cognition*, *1*, 83–95.

French, R. M., Mareschal, D., Mermillod, M., & Quinn, P. C. (2004). The role of bottom-up processing in perceptual categorization by 3- to 4-month-old infants: simulations and data. *Journal of Experimental Psychology: General*, *133*, 382–397.

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. New York, NY: Oxford University Press.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *8*(9), 404–409.

Gelman, S. A., & Davidson, N. S. (2013). Conceptual influences on category-based induction. *Cognitive Psychology*, *66*, 327–353.

Gelman, S. A., & Kalish, C. W. (2007). Conceptual development. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of Child Psychology* (Vol. 2, pp. 687–733). Hoboken, NJ: Wiley.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory*, *4*, 161–178.

Gentner, D., & Kurtz, K. J. (2005). Learning and using relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the laboratory* (pp. 151–175). Washington, DC: American Psychological Association.

Ghetti, S., & Lee, J. (2010). Children's episodic memory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 365–373.

Gliga, T., Volein, A., & Csibra, G. (2010). Verbal labels modulate perceptual object processing in 1-year-old children. *Journal of Cognitive Neuroscience*, *22*, 2781–2789.

Gliozzi, V., Mayor, J., Hu, J.-F., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: a neurocomputational approach. *Cognitive Science*, *33*, 709–738.

Goldstone, R. L. (1994). The role of similarity in categorization: providing a groundwork. *Cognition*, *52*, 125–157.

Goldstone, R. L., Kersten, A., & Carvalho, P. F. (2017). Categories and concepts. In J. Wixted (Ed.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed.). Hoboken, NJ: Wiley.

Greene, T. R. (1994). What kindergartners know about class inclusion hierarchies. *Journal of Experimental Child Psychology*, *57*, 72–88.

Gutheil, G., & Gelman, S. A. (1997). Children's use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology*, *64*(2), 159–174.

Halford, G. S., Andrews, G., & Jensen, I. (2002). Integration of category induction and hierarchical classification: one paradigm at two levels of complexity. *Journal of Cognition and Development*, *3*, 143–177.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: a new hypothesis. *Trends in Cognitive Sciences*, *11*, 236–242.

Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: what makes the difference? *Cognition*, *112*, 105–119.

Hanania, R., & Smith, L. B. (2009). Selective attention and attention switching: towards a unified developmental approach. *Developmental Science*, *13*, 622–635.

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition*, *50*, 171–188.

Hayes, B. K., & Heit, E. (2004). Why learning and development can lead to poorer recognition memory. *Trends in Cognitive Sciences*, *8*, 337–339.

Hayes, B. K., & Thompson, S. P. (2007). Causal relations and feature similarity in children's inductive reasoning. *Journal of Experimental Psychology: General*, *136*, 470–484.

Hayne, H. (1996). Categorization in infancy. *Advances in Infancy Research*, *10*, 79–120.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Categorical structure among shared features in networks of early-learned nouns. *Cognition*, *112*, 381–396.

Homa, D., Rhoades, D., & Chambliss, D. (1979). Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 11–23.

Inagaki, K., & Hatano, G. (2002). *Young Children's Naive Thinking about Biological World*. New York, NY: Psychology Press.

Inhelder, B., & Piaget, J. (1964). *The Early Growth of Logic in the Child: Classification and Seriation*. New York, NY: Harper & Row.

James, W. (1983). *The Principles of Psychology*. Cambridge, MA: Harvard University Press. (Originally published 1890.)

Japkowicz, N. (2001). Supervised versus unsupervised binary-learning by feedforward neural networks. *Machine Learning*, *42*, 97–122.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, *88*, 197–227.

Kemp, C., Shafto, P., & Tenenbaum, J. B. (2012). An integrated account of generalization across objects and features. *Cognitive Psychology*, *64*, 35–73.

Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General*, *137*, 52–72.

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*, 555–583.

Kripke, S. (1972). Naming and necessity. In D. Davidson & G. Harman (Eds.), *Semantics of Natural Language* (pp. 253–355). Dordrecht, the Netherlands: Reider.

Lazareva, O. F., & Wasserman, E. A. (2008). Categories and concepts in animals. In J. Byrne (Ed. in chief), R. Menzel (Ed.), *Learning and Memory: A Comprehensive Reference: Vol. 1. Learning Theory and Behavior* (pp. 197–226). Oxford, UK: Elsevier.

López, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development*, *63*, 1070–1090.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: redundant labels facilitate learning of novel categories. *Psychological Science*, *18*, 1077–1083.

Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review*, *99*, 587–587.

Mandler, J. M. (1999). Seeing is not the same as thinking: commentary on "Making sense of infant categorization." *Developmental Review*, *19*, 297–306.

Mandler, J. M., & Bauer, P. J. (1988). The cradle of categorization: is the basic level basic? *Cognitive Development*, *3*, 247–264.

Mandler, J. M., & McDonough, L. (1996). Drinking and driving don't mix: inductive generalization in infancy. *Cognition*, *59*, 307–335.

Mandler, J. M., & McDonough, L. (1998). Studies in inductive inference in infancy. *Cognitive Psychology*, *37*, 60–96.

Marks, W. (1991). Effects of encoding the perceptual features of pictures on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 566–577.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York, NY: Henry Holt.

Mather, E., & Plunkett, K. (2011). Same items, different order: effects of temporal variability on infant categorization. *Cognition*, *119*, 438–447.

Mayor, J., & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological Review*, *117*, 1–31.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 310–322.

McDonough, L., & Mandler, J. M. (1998). Inductive generalization in 9- and 11-month-olds. *Developmental Science*, *1*, 227–232.

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, *6*, 203–229.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, *44*, 1469–1481.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Merriman, W. E., Schuster, J. M., & Hager, L. (1991). Are names ever mapped onto preexisting categories? *Journal of Experimental Psychology: General*, *120*, 288–300.

Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*. New York, NY: Cambridge University Press.

Müller, U., & Overton, W. F. (1998). How to grow a baby: a reevaluation of image-schema and Piagetian action approaches to representation. *Human Development*, *41*, 71–111.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. *Psychological Review*, *81*, 267–285.

Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development of episodic and autobiographical memory. A cognitive neuroscience perspective. *Advances in Child Development and Behaviour*, *35*, 40–89.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 700–708.

Oakes, L. M., Coppage, D. J., & Dingel, A. (1997). By land or by sea: the role of perceptual similarity in infants' categorization of animals. *Developmental Psychology*, *33*, 396–407.

Opfer, J. E., & Siegler, R. S. (2004). Revisiting preschoolers' living things concept: a microgenetic analysis of conceptual change in basic biology. *Cognitive Psychology*, *49*, 301–332.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*, 185–200.

Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: a life-span overview. *Acta Psychologica*, *86*, 227–272.

Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, *106*, 665–681.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, *12*, 410–430.

Quine, W. V. (1960). *Word and object*. Cambridge, MA: MIT Press.

Quinn, P. C. (1987). The categorical representation of visual pattern information by young infants. *Cognition*, *27*, 145–179.

Quinn, P. C. (2002a). Beyond prototypes: Asymmetries in infant categorization and what they teach us about the mechanisms guiding early knowledge acquisition. In R. Kail & H. Reese (Eds.), *Advances in Child Development and Behavior* (Vol. 29, pp. 161–193). Oxford, UK: Elsevier.

Quinn, P. C. (2002b). Category representation in young infants. *Current Directions in Psychological Science*, *11*, 66–70.

Quinn, P. C. (2011). Born to categorize. In U. Goswami (Ed.), *Blackwell Handbook of Childhood Cognitive Development* (pp. 129–152). Oxford, UK: Blackwell.

Quinn, P. C., & Bhatt, R. S. (2006). Are some Gestalt principles deployed more readily than others during early development? The case of lightness versus form similarity. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1221–1230.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, *22*, 463–463.

Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy*, *1*, 31–46.

Rabi, R., & Minda, J. P. (2014). Rule-based category learning in children: the role of age and executive functioning. *PloS One*, *9*, e85316. http://doi.org/10.1371/journal.pone.0085316

Rakison, D. H., & Butterworth, G. E. (1998). Infants' use of object parts in early categorization. *Developmental Psychology*, *34*, 49–62.

Rhodes, M., Gelman, S. A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching and discovery. *Developmental Science*, *13*(3), 421–429.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review*, *7*, 158–162.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*, 1–20.

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*, 1387–1401.

Robinson, C. W., & Sloutsky, V. M. (2007a). Linguistic labels and categorization in infancy: do labels facilitate or hinder? *Infancy*, *11*, 233–253.

Robinson, C. W., & Sloutsky, V. M. (2007b). Visual processing speed: effects of auditory input on visual processing. *Developmental Science*, *10*, 734–740.

Robinson, C. W., & Sloutsky, V. M. (2010). Effects of multimodal presentation and stimulus familiarity on auditory and visual processing. *Journal of Experimental Child Psychology*, *107*, 351–358.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological Review*, *111*, 205–235.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, MA: MIT Press.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General*, *133*, 63–82.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel Distributed Processing* (pp. 318–362). Cambridge, MA: MIT Press.

Rumelhart, D. E., & Todd, P. M. (1993). Learning and connectionist representations. In E. D. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (pp. 3–30). Cambridge, MA: MIT Press.

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: an alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*, 94–104.

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition*, *73*, 1–33.

Scribner, S., & Cole, M. (1973). Cognitive consequences of formal and informal education. *Science*, *182*, 553–559.

Shipley, E. F. (1993). Categories, hierarchies, and induction. *Psychology of Learning and Motivation*, *30*, 265–301.

Siegler, R., & Svetina, M. (2006). What leads children to adopt new strategies? *Child Development*, *77*, 997–1015.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology*, *25*, 231–280.

Sloutsky, V. M. (2010). From perceptual categories to concepts: what develops? *Cognitive Science*, *34*, 1244–1286.

Sloutsky, V. M., Deng, W. S., Fisher, A. V., & Kloos, H. (2015). Conceptual influences on induction: a case for a late onset. *Cognitive Psychology*, *82*, 1–31.

Sloutsky, V. M., & Fisher, A. V. (2004a). Induction and categorization in young children: a similarity-based model. *Journal of Experimental Psychology: General*, *133*, 166–188.

Sloutsky, V. M., & Fisher, A. V. (2004b). When development and learning decrease memory: evidence against category-based induction in children. *Psychological Science*, *15*, 553–558.

Sloutsky, V. M., & Fisher, A. V. (2012). Linguistic labels: Conceptual markers or object features? *Journal of Experimental Child Psychology*, *111*, 65–86.

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: appearance similarity versus kind information in early induction. *Psychological Science*, *18*, 179–185.

Sloutsky, V. M., & Lo, Y.-F. (1999). How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental Psychology*, *35*, 1478–1492.

Sloutsky, V. M., Lo, Y.-F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*, *72*, 1695–1709.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*, 822–833.

Sloutsky, V., & Robinson, C. (2008). The role of words and sounds in infants' visual processing: from overshadowing to attentional tuning. *Cognitive Science*, *32*, 342–365.

Sloutsky, V. M., & Robinson, C. W. (2013). Redundancy matters: flexible learning of multiple contingencies in infants. *Cognition*, *126*, 156–164.

Smith, E. E., & Medin, D. L. (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: the early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1411–1436.

Smith, J. D., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 3–27.

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review*, *96*, 125–144.

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In B. Burnes (Ed.), *Percepts, Concepts, and Categories: the Representation and Processing of Information* (pp. 233–272). Amsterdam, the Netherlands: Elsevier.

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568.

Soja, N. N. (1992). Inferences about the meanings of nouns: the relationship between perception and syntax. *Cognitive Development*, *7*, 29–45.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition*, *38*, 179–211.

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: short-cutting abstraction in children's object categorizations. *Cognition*, *108*, 626–638.

Spencer, J. P., Blumberg, M. S., McMurray, B., Robinson, S. R., Samuelson, L. K., & Tomblin, J. B. (2009). Short arms and talking eggs: why we should no longer abide the nativist–empiricist debate. *Child Development Perspectives*, *3*, 79–87.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41–78.

Strauss, M. S. (1979). Abstraction of prototypical information by adults and 10-month-old infants. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 618–632.

Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: effects of retention interval and level of processing. *Memory and Cognition*, *29*, 424–432.

Tulving, E. (1972). *Organization of Memory*. Cambridge, MA: MIT Press.

Vygotsky, L. S. (1934). *Thought and Language*. Cambridge, MA: MIT Press. (Originally published 1934.)

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635–657.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Waxman, S. [R.], & Booth, A. (2003). The origins and evolution of links between word learning and conceptual organization: new evidence from 11-month-olds. *Developmental Science*, *6*, 128–135.

Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development*, *72*, 1653–1673.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, *34*, 1289–1309.

Wills, A. J., & Pothos, E. M. (2012). On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin*, *138*, 102–125.

Winer, G. A. (1980). Class-inclusion reasoning in children: a review of the empirical literature. *Child Development*, *51*, 309–328.

Wittgenstein, L. (2010). *Philosophical Investigations*. Hoboken, NJ: Wiley. (Originally published 1953.)

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, *85*, 223–250.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*, 245.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language*, *39*, 124–148.

Yamauchi, T., & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 776–795.

Younger, B. [A.]. (1990). Infant categorization: memory for category-level and specific item information. *Journal of Experimental Child Psychology*, *50*, 131–155.

Younger, B. A., & Cohen, L. B. (1985). How infants form categories. *Psychology of Learning and Motivation*, *19*, 211–247.

CHAPTER 3

Language Acquisition

JEFFREY LIDZ AND LAUREL PERKINS

INTRODUCTION

Language is about as close to magic as we can get. We push air through our lungs, vibrate our vocal cords, and move our mouths, and as a result, we can make the people around us become aware of past events, understand our thoughts or plans, perform actions, or come to have new beliefs. This magic is made possible by the shared cognitive systems, or grammars, of speakers and listeners. This shared grammar represents the sounds that make up the morphemes and words that bear meaning, and the rules of syntax that combine words into phrases and sentences that convey meaning. The study of language acquisition aims to uncover how this shared cognitive system arises within the mind of a human child. How does a child exposed to the vibrations of air caused by our utterances come to build a cognitive system for producing and understanding an unlimited number of sentences?

Answering this question requires a broad understanding of the kinds of tools that children use to solve the language learning problem. These tools include resources coming from extralinguistic cognition and from domain-specific biases that define grammatical knowledge for any language. Because language learners by necessity learn the language of their environment, a major contributor to language acquisition is likely to be the ability to track statistical information in the environment and to make use of patterns that are revealed in this information. This ability may be aided by other kinds of extralinguistic cognition, such as the perceptual capacities that shape how sounds are perceived as language or the conceptual capacities that undergird the meanings of words and sentences. Children must also use linguistic information in acquiring the grammar of their language. Partial knowledge in one domain of language may make available new resources for representing and identifying aspects of grammar in another domain of language. Similarly, architectural constraints on possible grammatical structures may also play a key role in shaping how children map their experience with language onto a grammatical system, essentially guiding them to look for certain kinds of information in their experience. Finally, because language is used predominantly as a tool for communication, understanding other people's goals and intentions will play a significant role in helping children to identify why people say the things they do, which in turn may contribute to their ability to identify the meanings of sentences.

In what follows, we consider how children identify the grammatical system that supports the ability to produce and understand new sentences, considering phonology, lexicon, syntax and semantics. In each case, we try to identify the independent contributions of experience, domain-specific biases, prior knowledge and extralinguistic cognition in shaping how a grammar grows inside the mind of a child.

PHONOLOGY

Perhaps the first task that learners must solve in acquiring a language is to identify its phonology, that is, the sound system of the language. Children must learn which acoustic variations in the speech they hear convey differences in meaning—that is, which ones come from the set of sound categories in their language, or phonemes, that speakers can combine into different words. Children must also learn the allophonic rules in their language that produce systematic variation within a sound category, so that a particular phoneme is pronounced differently depending on where it appears in a word. Learning phonology takes place together with word segmentation, the task of identifying word boundaries in a continuous speech stream. Children's abilities to track the statistical distributions of sounds and syllables in their input, combined with their developing knowledge of the rules in their language that govern those distributions, allow them to solve these two problems in tandem.

Phonemes and Rules

A phonology consists of two parts: a phonemic inventory, the set of sounds that are contrastive in the language, and a rule system, the system determining the linguistic environments in which particular sounds can and cannot occur.

A phonology is importantly different from the phonetics, which encompasses the articulatory processes involved in producing speech sounds and the acoustic properties that these sounds have. A phonology instead defines the distinctions that lead to meaningful differences in words and those that do not. For example, the [p] that occurs in the word [pɪt] is articulated differently from the one in the word [spɪt]. There is a longer delay between the release of the lip closure and onset of voicing associated with the vowel in the first word than in the second. The first [p] is aspirated, and the second is not. And this articulatory difference is reflected in the acoustics. But no words in English differ minimally in terms of this delay. This phonetic distinction is not contrastive and hence is not represented as a difference in phonological inventory of English speakers. The two distinct sounds are categorized as the same from the perspective of the phonology, just like a dachshund and a Great Dane are both categorized as dogs, despite their physical differences. Sounds that are both physically and psychologically distinct, such as [p] and [b], are contrastive, in that words that differ in these sounds also differ in meaning, as in [bɪt] versus [pɪt]. Importantly, not every phonetic distinction has a corresponding phonemic distinction. And languages differ with respect to which phonetic distinctions are treated as the basis for phonological categories and which are not.

The second component of a phonology concerns the rules governing the distributions of sounds in the language. Keeping to our example, the two [p]s just described are in complementary distribution—they cannot occur in the same word environments. The aspirated [p] occurs only when it is the first segment in a stressed syllable. The unaspirated [p] occurs in all other environments. So, we can say that these two sounds are related by rule to a single underlying category. The single phonemic category /p/ has two phonetic realizations, or allophones, determined by the phonological context.

The phonemic inventory and distributional rules vary from language to language and so must be learned. This learning process consists in identifying the underlying categories and determining the rules that govern the choice of allophones in different contexts.

Learning Phonemic Categories

In order to assess how children acquire the underlying categories, it is important to first understand how they perceive speech prior to acquiring these categories and then to ask how they use their experience to identify them.

Adult listeners demonstrate categorical perception: Although they can discriminate small differences within a category, perceptual discrimination is enhanced at category boundaries (Liberman, 1957; McMurray, Tanenhaus, & Aslin, 2002). When presented with computer-generated stimuli that either fall within a single category or cross a category boundary, adults discriminate better when the pair of sounds crosses a boundary than when it falls within a boundary, even if the size of the acoustic difference is identical. Infants, like adults, show enhanced discrimination of acoustic-phonetic differences that cross category boundaries (Dehaene-Lambertz & Dehaene, 1994; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Werker & Lalonde, 1988). This is true even for categories that are not represented in the language in the child's environment (Eimas et al., 1971; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1984). These sensitivities are also shared across species, suggesting that they reflect basic perceptual processes and are not strictly linguistic in nature (Kuhl & Miller, 1975; Mesgarani, David, Fritz, & Shamma, 2008; Ramus, Hauser, Miller, Morris, & Mehler, 2000).

Because phonological categories vary across languages, infants must learn which distinctions are meaningful in their language. For example, Hindi contains a contrast between an alveolar [d] and a retroflex [D] that is not represented in English (although the natural variability in English /d/ sometimes includes retroflex pronunciations, as in sequences like *our doll*). Infants at 6 to 8 months of age are able to discriminate these sounds, unlike English speaking adults, though this ability declines by around the first birthday (Werker & Tees, 1984). This widely replicated pattern of broad sensitivity in young infants followed by language-specific discrimination in older infants and adults indicates that the development of phonological categories involves maintenance of initial auditory sensitivities rather than the creation of new categories from a perceptually neutral acoustic space (Kuhl et al., 2006; Narayan, Werker, & Beddor, 2010; Polka & Werker, 1994). The initial perceptual sensitivities of infants are maintained or sharpened as a function of experience (Kuhl et al., 2006; Maye, Weiss, & Aslin, 2008; Narayan et al., 2010), but there is no evidence that brand new phoneme categories can be induced solely from language listening experience.

The loss of the "universal listener" abilities involved in the identification of phoneme categories does not involve pruning perceptual abilities, however. The auditory system retains its categorical perception for nonnative speech sounds if the stimuli are not presented as speech (Werker & Tees, 1984). Instead, learning phoneme categories involves a functional reorganization, whereby initial perceptual distinctions get recoded as linguistic distinctions (Nazzi, Bertoncini, & Mehler, 1998; Werker, 1995). These linguistic distinctions may be identified from the distributional characteristics of the speech in the environment. Phonemic distinctions will be expressed through distributions that highlight the existence of two categories (e.g., many instances of [d] and [D] with fewer tokens falling in the space between these extremes), whereas variation of tokens within a category will be expressed through a more uniform distribution.

Maye, Werker, and Gerken (2002) showed that infants can track such frequency information and use it to change their phonetic category boundaries. Six- to 8-month-old infants were presented with stimuli from an 8-step voicing continuum from [da] to [ta]. In one condition, many of the items fell along the extreme ends of the continuum, with few in the middle, yielding a bimodal distribution of tokens. In the other condition, the most frequent items fell in the middle of the continuum, yielding a unimodal distribution of tokens. After 2 to 3 minutes of familiarization, infants were tested on their ability to discriminate the endpoints of the continuum. Those infants who were familiarized to the bimodal distribution showed better discrimination than those familiarized to the unimodal distribution.

Another source of information for building phonetic categories comes from their phonological environments. Feldman, Myers, White, Griffiths, and Morgan (2013) exposed 6- to 8-month-old infants to a uniform distribution of vowels from a continuum between [a] ('ah') and [ɔ] ('aw'). Half of the babies heard each of the sounds in distinct word forms (i.e., *gutah* versus *litaw*). The other half of the babies heard both vowels in both word forms (i.e., *gutah*, *gutaw*, *litah*, *litaw*). After being familiarized to these words, infants were then tested to see if they could distinguish between alternating pairs of syllables (*tah* versus *taw*) with vowels drawn from the ends of the continuum versus repetitions of a single syllable with a vowel drawn from the center of the continuum. Only those infants who were familiarized to distinct word forms were able to discriminate the alternating syllables from the nonalternating ones. Thus, the occurrence of sounds in distinct phonological environments provides evidence for learners about the identity of the sounds.

Learning Allophonic Rules

The categories that are built for linguistic representation feed forward into the learning of allophonic rules, the rules that govern alternations of sounds from within a single category. Seidl, Cristià, Bernard, and Onishi (2009) familiarized English- and French-learning infants with a pattern that linked the choice of a stop [t] versus a fricative [s] to the quality of the preceding vowel. Specifically, infants heard syllables in which nasal vowels (produced with airflow through the nasal cavity) were followed only by fricatives, and oral vowels (produced with airflow only through the oral cavity) were followed only by stops. English-learning 4-month-olds were able to learn this dependency. In French, nasal vowels contrast with oral vowels, so French-learning 11-month-olds were also able to learn this rule. However, English-learning 11-month-olds, who are acquiring a language in which the oral-nasal difference is allophonic, were not able to learn the rule. Because the English-speaking infants had acquired a single category containing both oral and nasal vowels, they were unable to learn a rule that depended on nasality. The very same sounds function differently in the mental representation of speech by the end of the first year of life, and these representations feed forward for subsequent learning.

Similarly, Onishi, Chambers, and colleagues (Chambers, Onishi, & Fisher, 2003; Onishi, Chambers, & Fisher, 2002) taught 16-month-old infants two kinds of phonotactic constraints, which are language-specific restrictions on which sequences of sounds are possible and where in a syllable certain sounds can occur. Infants learned a simple positional regularity in which /b/ was allowed only as the first sound in a syllable, and a context-dependent regularity in which /b/ occurred after /ae/ but not after /i/. Infants were able to learn both kinds of regularity. Importantly, conditioning distributions of consonants on the speaker's voice or identity in a third study did not induce learning. That is, children were able to learn rules conditioning sound distributions on linguistic information like word environments, but not on speaker identity, a language-external factor. Thus, these experiments not only demonstrate infants' rich abilities to learn novel distributional constraints on allophones, but they also indicate that such learning is restricted to the kinds of regularities that languages regularly encode.

Word Segmentation

Acquisition of the phonological inventory of a language takes place concurrently with the acquisition of word-segmentation abilities. (See Nazzi et al., 2016, for review.) Segmentation of word forms from the speech stream also plays a critical role in the acquisition of words as pairings of form and meaning. Word segmentation abilities at earlier ages are predictive of vocabulary size at later ages (Newman, Ratner, Jusczyk, Jusczyk, & Dow, 2006), and newly segmented words are easier for 17-month-old infants to link to a meaning than are wholly novel words (Graf Estes, Evans, Alibali, & Saffran, 2007).

Jusczyk and Aslin (1995) showed that word segmentation abilities first develop between 6 and 8 months of age. They familiarized 7.5-month-old English-learning infants to two monosyllabic words (*cup* and *dog* or *bike* and *feet*). Infants then heard four passages, each containing six repetitions of one of the four words. Infants showed a preference for the passages containing the familiarized words, indicating that they had recognized those words. This recognition implies that they were able to segment the familiarized words from the passages. This result failed to extend to 6-month-old infants.

Six-month-olds can segment words under some circumstances, however. Unlike Jusczyk and Aslin (1995), Bortfeld, Morgan, Golinkoff, and Rathbun (2005) showed that 6-month-old infants could segment unfamiliar words from a passage if these words were preceded by highly familiar words, such as the infant's name or the word "mommy." (See Brent & Cartwright, 1996, for a computational model of word segmentation based on familiar words.)

Infants as young as 7.5 months of age can use the rhythmic units of their language to segment words. In English, a majority of words are stressed on the first syllable (Cassidy & Kelly, 1991; Cutler & Carter, 1987). Correspondingly, English-learning infants between 6 and 9 months of age show a preference for stress-initial words over stress-final words (Jusczyk, Cutler & Redanz 1993). This result does not derive from a general perceptual preference for stress-initial words; French-learning infants between 4 and 6 months of age show the opposite preference (Friederici, Friedrich, & Christophe, 2007). These preferences also contribute to segmentation. Jusczyk, Houston, and Newsome (1999) showed that English-learning 7.5-month-olds successfully segment trochaic (strong-weak) words such as "doctor" from a passage, but that they missegment iambic (weak-strong) words such as "guitar." By 10.5 months, infants are able to successfully segment the iambic words as well.

By 8 months of age, infants may also be able to use order of syllables within words as a cue to the location of word boundaries. Saffran, Aslin, and Newport (1996) showed that 8-month-old infants use the transitional probability between two syllables (i.e., the probability of two syllables occurring together) as a cue to word boundaries, based on the assumption that two syllables that frequently co-occur are part of the same word (Brent & Cartwright, 1996). These authors familiarized infants to 2 minutes of continuous speech made up of randomly concatenated sequences of four trisyllabic "words," such as "pabiku, todabu." They then tested infants' listening preferences to these words as compared to sequences of syllables that had occurred in the familiarization, but were taken from different words, so that they exhibited a lower transitional probability. Infants listened longer to the "words," suggesting that they use the transitional probabilities as evidence for where the word boundaries occur in this artificial language. Eight-month-olds are unable to use transitional probabilities to segment words of varying length (E. K. Johnson & Tyler, 2010; Mersad & Nazzi, 2012); however, they could do so if one of the words was a highly familiar word. This finding suggests that infants can combine multiple sources of information in word segmentation.

By 9 months, infants can use phonotactic properties of their language to help segment words from the speech stream. By 6 months, infants show a preference for sequences of sounds that are possible in their language over sequences that are impossible (Jusczyk, Luce, & Charles-Luce, 1994). And by 9 months, they demonstrate better segmentation of words with high between-word phonotactic probabilities at their edges (e.g., [zt]) than words with high within-word probabilities (Mattys & Jusczyk, 2001).

Between 10 and 12 months of age, infants can use the prosody of their language, or its stress and intonation patterns, as a cue to word segmentation. Children are sensitive to breaks between prosodic units at extremely early ages. Newborns can tell the difference between stimuli with and without prosodic breaks (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Christophe, Mehler, & Sebastián-Gallés, 2001). By the age of 9 months, infants have developed knowledge about prosodic phrases in their language: They react when a prosodic phrase is disrupted by the insertion of a break (Gerken, Jusczyk, & Mandel, 1994; Jusczyk et al., 1992). And 10- and 12-month-olds can use those prosodic breaks to constrain word identification (Christophe, Millotte, Bernal, & Lidz, 2008; Gout, Christophe, & Morgan, 2004; Millotte et al., 2010). Gout et al. (2004) conditioned infants to turn their heads when they heard the word "paper." Then, they exposed the infants to sentences in which that same sequence of syllables occurred within a phonological phrase (e.g., *[the scandalous* ***paper****] [sways him] [to tell the truth]*), or across a phonological phrase boundary (e.g., [*the outstanding* ***pay****]* *[****per****suades him] [to go to France]*). They found that these infants turned their heads more often when the syllables occurred within a phonological phrase than when it spanned a phonological phrase boundary, indicating that the presence of a phrase boundary disrupted their ability to recognize the word "paper." These results suggest that children use phonological phrasing as a cue for word segmentation.

In sum, infants' ability to segment words from the speech stream undergoes significant development between 6 and 12 months of age. At first, they can segment only highly frequent words. Over time, infants develop word segmentation strategies based in the rhythmic properties of words, the statistical properties of the syllable transitions, and the phonotactic and prosodic features of their language.

Summary

In acquiring the phonology of their language, children organize the acoustic information in the speech signal into phonemic categories and infer the allophonic rules that specify how sounds from one category systematically vary depending on surrounding sounds or their position in a word. Properties of children's extralinguistic auditory system allows them to perceive all sound contrasts that languages might make use of, but based on the statistical distribution of the sounds they hear, children eventually form language-specific representations that encode only the meaningful contrasts in their language. Simultaneously, children learn to segment the continuous speech stream into words, aided by their statistical sensitivity to syllable distributions as well as their knowledge of the rhythmic properties, prosody, and phonotactics of their language. Next we look in more depth at how children learn what those words mean and how to use them.

LEXICON

Consider what you, as a proficient adult speaker of a language, know about your lexicon, or the set of words in your language. You know their phonological forms; you can pronounce words like a native speaker and identify them in the speech of others around you. In the previous section, we discussed how children learn the phonology of their language and solve the problem of word segmentation. You also know what words mean; you can map from those phonological forms to the specific concepts they pick out. These mappings from sound to meaning are arbitrary and language-specific. If you speak English, the phonological form /fi/ means a cost or a charge (*fee*), but if you speak French that same phonological form means a girl or daughter (*fille*).

But before we consider how children learn the mappings from sounds to meanings in their language, we consider another part of your knowledge about words: how to use them in sentences. As an adult speaker of English, you know that the word "arrive" can be used after a helping verb like "will": You can say *Elliott will arrive*. You also know that the word "arrival" can be used after an article like "the": You can say *The arrival of Elliott was unexpected*. Furthermore, you know which environments these words cannot occur in: You cannot say **Elliott will arrival* or **The arrive of Elliott was unexpected.* (The asterisk indicates that a string of words is unacceptable in a particular language.) Part of your knowledge about words includes features that we call grammatical categories, which determine their distributions in sentences. Even though "arrive" and "arrival" have similar meanings, their different grammatical categories (verb versus noun) lead to different sentence distributions. Learning which words belong to which grammatical categories is yet another problem that children need to solve when learning their lexicon.

Grammatical Categories

Grammatical categories—such as noun, verb, and adjective—are names for the features that determine which syntactic environments lexical items can appear in. When we say *arrival* is a noun, we mean that the "noun" feature allows *arrival* to occur after articles but not after helping verbs, for example. When we say *arrive* is a verb, we mean the "verb" feature allows *arrive* to occur after helping verbs but not after articles.

These grammatical categories sometimes correlate with semantic categories, but there are many exceptions: we're often told that a noun is a "person, place, or thing" and that verbs describe actions, but the verb *believe* does not really describe an action, whereas the noun *destruction* does.

Grammatical categories come in two flavors: lexical and functional. Lexical categories include the familiar categories of noun, verb, and adjective—these are what we might call "content words," and they are also open class, meaning that we can easily coin new words that fall into these categories. Functional categories are closed class, meaning that it is hard or impossible to coin new words in these categories, and they contain less referential content. Some functional categories include determiners, such as *the*, *a*, *some*, *most*; pronouns, such as *I*, *you*, *he*, *she*, *it*; modals and auxiliaries ("helping verbs"), such as *have*, *be*, *may*, *will*, *can*; and morphemes (pieces of words) that signal tense and agreement, such as past tense *-ed*, present progressive *-ing*, and plural *-s*. Functional categories frequently signal when specific lexical categories are upcoming; for example, determiners are signals for nouns. These signals might be useful information in children's learning processes.

Using Distributional Information to Categorize Words

Because a word's grammatical category determines its distribution in sentences, children may be able to use that distributional information as a signal for the grammatical categories of new words. Computational simulations have probed the extent to which distributional regularities in speech to children can support word categorization. Many of these simulations have achieved fairly high success in categorizing words into grammatical categories based solely on these patterns in how words cluster together and which words tend to occur next to each other (Cartwright & Brent, 1997; Mintz, 2003; Mintz, Newport, & Bever, 1995, 2002; Redington & Chater, 1998; Redington, Chater, & Finch, 1998). For example, Mintz (2003) used an algorithm that clustered words based on similarities in their immediately preceding and following sentence environments ("frames"). This algorithm was reliably able to separate nouns from verbs based only on the information in these frames. These types of models show that there is a distributional signal for the grammatical categories of words in speech to children. Furthermore, many experimental studies have found that children are skilled at detecting and using that signal.

From extremely early ages, children appear sensitive to the differences between function words and content words, which tend to have different acoustic and phonological properties cross-linguistically. Across languages, function words are often unstressed, shorter than content words, have reduced vowels, and appear at prosodic boundaries (e.g., Monaghan, Chater, & Christiansen, 2005; Shi, Morgan, & Allopenna, 1998). Even newborns demonstrate sensitivity to these differences. In a study by Shi, Werker, and Morgan (1999), newborns heard repetitions of English words selected from an audio recording of natural maternal speech. Infants' attention to these audio stimuli was tested using a procedure called high-amplitude sucking, which measures infants' sucking strength and rate on pressure-sensitive pacifier. Infants learn that they can control the presentation of an audio stimulus by sucking harder, and the researchers measure how the rate of these high-amplitude sucks declines over time as infants lose attention. Once this rate declines to a certain threshold, infants are considered to be "habituated" to the stimulus, and a new test stimulus is played. If infants consider this new stimulus different from the previous one, they should recover attention ("dishabituate") and therefore increase their rate of high-amplitude sucks. Shi et al. habituated infants to a list of either content words or function words and then tested them on new words from the same category or the opposite category. Infants who were habituated to content words recovered attention and increased their sucking rate when they heard function words, and vice versa, but did not recover attention when they heard new content words. It therefore appears that newborns are able to discriminate the phonological differences between function and content words. This ability may enable infants to begin categorizing words into functional and lexical categories from the earliest stages of language acquisition.

Learning the specific phonological forms of function words in the infant's target language takes place over the first year of life. Infants are able to segment function words in their own language by the age of 6 months (Höhle & Weissenborn, 2003; Shi, Marquis, et al., 2006), and differentiate real function words from phonologically similar nonsense function words between the ages of 8 and 11 months (Hallé, Durand, & de Boysson-Bardies, 2008; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008; Shi, Werker, & Cutler, 2006). Children at early stages of sentence production frequently omit function words in their own speech, but repeat sentence prompts with real and nonsense function words at different rates, indicating that they know the difference (Gerken, Landau, & Remez, 1990).

Once the forms of function words are learned, they become useful in learning other new words. Early on, they can serve as anchors in the speech stream: 8-month-olds can use known function words to segment new content words (Shi & Lepage, 2008). Older infants can use function words as a signal for specific lexical categories. For example, 14- to 16-month-olds who are familiarized with a nonsense word preceded by a determiner (e.g., *my kets*) react with surprise when the same nonsense word occurs in an environment in which nouns cannot occur, such as after an auxiliary (*will kets*) (Hicks, Maye, & Lidz, 2007; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Shi & Melançon, 2010). Infants also react with surprise when a nonsense word preceded by a modal (*will dak*) is later preceded by a determiner (*my dak*) (Hicks et al., 2007). This finding suggests that children use the determiner and auxiliary functional categories to identify the lexical category of an unknown word: Hearing a determiner tells them that this word is a noun and therefore should occur only in places where nouns can occur, and hearing an auxiliary tells them that this word is a verb and should occur only in places where verbs can occur.

Children deploy their knowledge of function words during online language comprehension to help identify known words. A study by Cauvet et al. (2014) trained 18-month-old French-learning children to respond to a target noun preceded by a determiner (e.g., *la balle*, "the ball") or a target verb preceded by a pronoun (*je mange*, "I eat"). At test, children recognized the target words more frequently when they were preceded by another word from the correct functional category—when the target noun was preceded by a determiner or when the target verb was preceded by a pronoun. Other studies have found that 2-year-olds show better and faster sentence comprehension when singular nouns are preceded by determiners than by ungrammatical or missing function words (Gerken & McIntosh, 1993; Kedar, Casasola, & Lust, 2006; Shipley, Smith, & Gleitman, 1969).

Furthermore, children can use functional categories to infer aspects of a content word's meaning. Even though grammatical categories do not correlate perfectly with semantic categories, some imperfect correlations do exist: For example, nouns tend to label object kinds, adjectives tend to label object properties, and verbs tend to label events. Children as young as 1 year old can use known function words to infer whether a novel word labels an object kind or property (Hall, Waxman, & Hurwitz, 1993; Mintz & Gleitman, 2002; Smith, Jones, & Landau, 1992; Taylor & Gelman, 1988; Waxman, 1999; Waxman & Booth, 2001; Waxman & Markow, 1998). Twelve-month-olds who hear an object labeled as *a blicket* will select another object of the same kind when asked for another blicket (Waxman & Markow, 1998). Thirteen-month-olds who hear a purple horse labeled as *a daxish one* will prefer to select a novel purple object over a different-colored horse (Waxman, 1999). This behavior suggests that 1-year-old infants can distinguish the distribution of nouns and adjectives based on co-occurring functional categories and use that knowledge to infer that a novel word in a noun context labels an object kind whereas a novel word in an adjective context labels an object property.

Slightly older infants are also able to use the presence of functional verbal morphology to identify that a novel word labels an event rather than an object. He and Lidz (2017) habituated 18-month-olds to a scene of a penguin spinning, labeled either by a novel word in a noun context (e.g., *It's a doke*) or in a verb context (*It's praching*). At test, children saw a scene of the penguin performing a different action, labeled by the same audio. Children dishabituated when they heard *It's praching* labeled that new scene but not when they heard *It's a doke*. These infants appear to have used the co-occuring functional categories to identify whether the novel word was a noun or verb and therefore what concept it should label. Infants who heard the novel word after a determiner identified the word as a noun and therefore an object name and were not surprised to hear this word label the same object performing a different action. By contrast, infants who heard the novel word with verbal morphology (*-ing*) identified the word as a verb and therefore an event name and were surprised to hear this word label a different action. Identifying the signals of a new word's grammatical category—its distributional context and co-occurring function words—allows children to both categorize and make inferences about the meaning of that word.

Bootstrapping from Prosody

In addition to distributional information, children's knowledge of the prosodic features of their language may feed their categorization of words. Recall that children are sensitive to prosodic breaks in their language from a very young age. If these breaks typically fall at the edges of phrases centered around certain grammatical categories, then children might be able to use them to identify those phrase boundaries and differentiate words of different grammatical categories. This process of using prosodic information to infer something about the syntactic properties of a phrase or clause is called prosodic bootstrapping (Christophe et al., 2008; de Carvalho, Dautriche, & Christophe, 2016; Gleitman, Gleitman, Landau, & Wanner, 1988; Gout et al., 2004; Gutman, Dautriche, Crabbé, & Christophe, 2015; Morgan, 1986; Morgan & Demuth, 1996; Morgan, Meier, & Newport, 1987; Morgan & Newport, 1981; Wanner & Gleitman, 1982).

A study by de Carvalho et al. (2016) found that French-speaking preschoolers can use the position of a prosodic break to identify the category of an ambiguous word. Four-year-olds were asked to complete a sentence fragment that contained a noun/verb homophone, such as *ferme*, which can either mean "farm" (a noun) or "close" (a verb). The category of the word was disambiguated by prosody. In the sentence [*la petite ferme*] [*est très jolie*] ("the little farm is very nice"), the prosodic break after *ferme* indicates that it is a noun; by contrast, in the sentence [*la petite*] [*ferme la fenêtre*] ("the little girl closes the window"), the prosodic break before *ferme* indicates that it is a verb. After hearing *la petite ferme,* 4-year-olds who heard a prosodic break before *ferme* provided completions indicating that they interpreted *ferme* as a verb, whereas children who heard no prosodic break interpreted the word as a noun. A similar result was found for 3-year-olds in a looking time experiment. These results suggest that preschoolers can exploit prosodic information in quite sophisticated ways: The prosodic breaks in a sentence allow them to identify which prosodic phrase contains an ambiguous word, and therefore to determine whether the word should be analyzed as a verb or a noun during online sentence comprehension.

Why would identifying the prosodic phrase containing a word be useful in identifying the category of that word? On one hypothesis, children's prosodic knowledge interacts with their knowledge about function words in their language. If function words tend to occur at the edges of prosodic phrases, then these words might help children categorize the co-occurring content words that the phrases are built around (Gutman et al., 2015; Morgan, 1986; Morgan & Demuth, 1996). For example, children might perceive that a string of words, such as *The cute little girl will dance*, contains two prosodic phrases, one starting with a determiner and one with a modal: [*The cute little girl*] [*will dance*]. If children know to pay attention to these function words at the edges of prosodic phrases and know that determiners signal nouns and modals signal verbs, then they might be able to label these phrases: The first is a noun phrase, and the second is a verb phrase. A computational model by Gutman et al. (2015) was able to differentiate noun phrases from verb phrases in child-directed speech with fairly high reliability by tracking the distribution of function words at the edges of prosodic phrases and by building off of a small "seed" of known object and action words.

Properties of Grammatical Categories

We have seen that statistical sensitivity to the distribution of words in the input, in conjunction with prosodic knowledge, can help children categorize words in their language. But children also use category information to infer properties of new words: for example, that words used in noun contexts label object kinds, words used in adjective contexts label object properties, and words used in verb contexts label events. This knowledge does not emerge straightforwardly from the distribution of these words in the input but rather from some knowledge about the types of meanings these categories can have. Where does this knowledge about the relation between grammatical categories and meanings come from?

It is possible that some knowledge about the properties of linguistic categories may be intrinsic to the language learning mechanism. Pinker (1984, 1989) was an early proponent of the hypothesis that the language learning mechanism contains knowledge of innate linking rules between meanings and the syntactic forms that can express those meanings. Therefore, understanding properties of the meaning of an utterance can help the learner infer syntactic properties of that utterance: a strategy called semantic bootstrapping. If the learner has innate knowledge that nouns label object kinds and verbs label events, then words that speakers seem to use to label object kinds must be nouns, and words that speakers use to label events must be verbs.

Conversely, if a group of words appear in a distribution that would indicate they are nouns, those words are likely to label object kinds; if a group of words appear in a distribution that would indicate that they are verbs, then those words are likely to label events.

We have already seen that children by the age of 18 months appear to know these linking rules. The 18-month-olds tested by He and Lidz (2017) understood that a consequence of being a noun meant that *doke* referred to an object, whereas a consequence of being a verb meant that *praching* referred to an event. However, this behavior does not necessarily prove that these linking rules are innate; it is possible that children could have learned the relations between grammatical categories and meanings by the time they are 18 months. Gutman et al.'s (2015) computational model was able to learn more general categories of nouns and verbs by tracking prosodic breaks and function words, when seeded with knowledge of words for a few common objects and actions. The authors proposed that children might learn the semantic properties of noun and verb categories by noticing that a few common words map onto perceptually salient categories such as concrete objects and causal actions. That is, if children are already likely to perceive their world in categories such as "object" and "action" and can learn that some words are used to label these categories, they might conclude that words having similar distributions are likely to label similar categories. Words that distribute like known words for objects are also going to label objects, and words that distribute like known words for actions are also going to label actions. If this hypothesis is correct, then the linking rules that children know by 18 months would be a consequence of the way they perceive the world in certain categories and the way they expect language to reflect those categories.

Further work is needed to determine whether the semantic properties of nouns, verbs, and adjectives are innately specified or learned through experience. However, there are other syntactic properties that follow as consequences of a word's grammatical category, many of which would be difficult or impossible to learn by observation. Pinker (1984) highlighted one important syntactic consequence of being a noun or a verb that may fall into the impossible-to-observe category. Both nouns and verbs can take full clauses as complements: You can say *Aaron claimed [that Bill saw Eva]* or *Aaron believed the claim [that Bill saw Eva]*. Furthermore, in the first sentence, you can question part of the embedded clause: *Who did Aaron claim [that Bill saw]?* However, you cannot question the exact same part of the embedded clause in the second sentence: **Who did Aaron believe the claim [that Bill saw]?* is not a possible question about the person Bill saw. English speakers can question parts of the clausal complements of verbs, but not nouns. This is a constraint that would be very difficult to learn by observation, because children are not able to observe which sentences in their language are not possible, only the ones that *are* possible. Furthermore, this constraint seems to hold cross-linguistically. For these reasons, Pinker and many others hypothesized that the constraint on questioning parts of the clausal complements of nouns comes from innately specified linguistic knowledge. If children have knowledge about the constraints on question formation that are obeyed by all human languages, all they will need to learn is whether a word is a noun or a verb, and they will know whether it is possible to question the clausal complements of that word.

The question of whether domain-specific knowledge or learning through experience is responsible for children's awareness of word properties also has been hotly debated for functional categories like determiners. Children display very early sensitivity to the presence of determiners in their input, but their early speech tends to omit determiners, leading some to wonder when children know that these words are part of the same grammatical category "determiner." Valian (1986) studied the speech of six 2-year-olds and found that when these children did use determiners, they used them in appropriate sentence environments and differentiated them from other prenominal categories, such as adjectives. Valian, Solt, and Stewart (2009) and Yang (2013) found that children use the category "determiner" productively from the earliest ages at which they start combining words: Once they begin producing multiple different determiners, they use these determiners interchangeably before nouns, indicating that they consider them members of the same category and know the distribution of this category. These results are in contrast to a claim that children at the relevant age lack the determiner category (Pine & Lieven, 1997; Pine & Martindale, 1996; Tomasello, 2000, 2003).

Valian and colleagues (2009) hypothesized that children are innately aware of the range of categories that languages can make use of, "determiner" being one of these categories, and therefore determiner acquisition involves mapping words in their language to this category. In order to provide strong support for the innateness hypothesis, it would be necessary to show that very young children have not only grouped determiners together based on their distribution in the input but are aware of what it means to be a determiner—that determiners have specific properties, which lead to specific constraints on their behavior.

Studies with older children have shown that they are sensitive to some interpretive consequences of being a determiner. In a study by Wellwood, Gagliardi, and Lidz (2016), 4-year-olds heard a novel word in a determiner context (e.g. *gleebest of the cows*), in an adjective context (*the gleebest cows*), or in a context where either a superlative determiner or adjective could occur (*the gleebest of the cows*). Children were asked to choose from a set of cards that showed multiple spotted cows. On some cards, most of the cows were by the barn but the spottiest cows were not; on other cards, the spottiest cows were by the barn but most of the cows were not. When children heard *gleebest* in a determiner context, such as *Gleebest of the cows are by the barn*, they preferred the cards where most cows were by the barn but the spottiest cows were not. When children heard *gleebest* in other contexts, they preferred the opposite cards. It appears that children assigned the novel word a quantity-based meaning only when it occurred in the determiner context; they assigned the word a quality-based meaning otherwise. These children were able to use the context in which the novel word occurred to categorize it as a determiner or adjective. Furthermore, they knew that only determiners, not adjectives, are restricted to having quantity-based interpretations.

In summary, children's knowledge about grammatical categories in their language goes beyond the distribution of these categories and includes information about other syntactic or interpretive properties of these categories. Children know that nouns label objects, adjectives label object properties, and verbs label events; they also know that determiners but not adjectives can have quantity interpretations. Furthermore, a word's grammatical category will have consequences on other dependents of that word in a sentence: It is possible to question part of a clause introduced by a verb but not by a noun. Some of these properties might be learned through observation—for instance, by observing that certain categories of words map onto perceptual categories such as "object" and "action." But many of these properties are difficult or impossible to observe, and yet seem to hold true cross-linguistically—such as constraints on question formation. In order for children to learn these constraints consistently in the face of very scarce evidence, it is likely that their learning process is guided in part by domain-specific linguistic knowledge, intrinsic to the language-learning mechanism.

Lexical Meanings

So far we have discussed how children learn the grammatical categories of words in their language as well as some semantic and syntactic properties associated with those categories. Now we consider how children learn the meanings of specific words in their language. Recall that a word's sound does not signal its meaning. The meaning "a black-and-white farm animal that produces milk" is encoded by the sequence of sounds [kaʊ] (*cow*) in English and [vaʃ] (*vache*) in French. The sound sequence [fi] means a cost or charge in English (*fee*) and a girl or daughter in French (*fille*). How do children learn these arbitrary, language-specific mappings between form and meaning? There are two possible signals: the situations in which the word is used and the word's syntactic properties. Here, we' discuss how statistical sensitivity, extralinguistic cognition, and prior linguistic knowledge may help children detect these signals and use them to draw inferences about word meanings.

Learning by Observation

One very old hypothesis about word learning, dating back to the philosopher John Locke (1690/1998), proposes that children can learn the meanings of words by observing what they are being used to label in the world. For example, English-speaking children hear the sequence of sounds [kaʊ] frequently in contexts where cows are present and learn that those sounds are used to label cows. This strategy has been called word-to-world mapping (Gleitman, 1990): A child learns the meaning of a word by observing the real-world contingencies of its use, or what possible referents for the word are present in the world when the word is being uttered.

Extralinguistic cognition could be very helpful in using this word-to-world mapping strategy, particularly if a child can figure out what in the world a speaker is referring to when using a particular word. Young children are adept at detecting some nonverbal cues that indicate what adults are referring to when they speak. Infants as young as 9 to 12 months old can follow a pointing finger and a speaker's eye gaze to locate what a speaker is attending to (see, e.g., Baldwin & Moses, 1996), and slightly older children use the speaker's eye gaze as a cue to the referent of a novel word (Baldwin, 1991, 1993). For example, the 18- and 19-month-olds in Baldwin's (1991, 1993) experiments checked what a speaker was attending to when they heard a novel word spoken and interpreted the object that the speaker was looking at as the referent of that word, even when another object was more salient.

But it is possible that cues such as eye gaze and the physical presence of the referent when the word is being uttered will be more reliable for learning certain types of words than others. Gillette, Gleitman, Gleitman, and Lederer (1999) conducted a simulation of word learning with adults in order to investigate this question: Does the extralinguistic context in which a noun or a verb is uttered provide enough information to infer its meaning, or is it more helpful for some words than for others? The experimenters presented adult participants with videos of mother–child interactions, in which the most common nouns and verbs uttered by the mother were indicated by a beep, and asked participants to guess what word the beep stood for. These adults were able to identify the correct noun 45% of the time based on the visual information alone but could identify the correct verb only 15% of the time. Later simulation studies, such as by Medina et al. (2011), found a similar result: In general, visual contexts seem to be more informative for identifying nouns than verbs. This asymmetry parallels the acquisition trajectories of nouns and verbs in many different languages: When children begin talking, they produce nouns almost exclusively, and verbs come later (Bates, Dale, & Thal, 1995; Caselli et al., 1995; Gentner, 1982). Perhaps this order of acquisition is related to how strongly extralinguistic information supports learning nouns, as opposed to verbs, by observation.

Statistical sensitivity may also help a learner infer the meaning of a word through observation. Many studies have found that children can also be quite good at fast mapping: learning the meaning of a word from a single presentation (Baldwin, 1993; Carey, 1978; Carey & Bartlett, 1978; Dollaghan, 1985; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Heibeck & Markman, 1987). But if the context is not informative about a word's meaning the first time a child hears it, the child may be able to track information about what the word is being used to label across many different exposures. This strategy is called cross-situational learning (e.g., Blythe, Smith, & Smith, 2010; Smith & Yu, 2008; Vouloumanos, 2008; Xu & Tenenbaum, 2007; Yu & Smith, 2007). For example, Smith & Yu (2008) presented 12-month-olds with pictures of geometric shapes paired with novel words. During each trial, two words were presented in the context of two shapes, such that the pairing of each word with its referent was ambiguous. However, each word–shape pairing was disambiguated over the course of 12 trials. Averaging across all of the trials, the authors found that these infants preferred to look at the correct referent for the majority of the novel words. It appears that these infants were able to map multiple labels to multiple objects by tracking how these labels were used across different trials, even though each single presentation of a word was ambiguous with regard to its referent.

Thus, although learning by observation can be difficult for certain word categories and in ambiguous contexts, it might be easier if learners can track statistical evidence for a word's meaning across many different contexts. But a more fundamental issue remains. The particular word that a speaker uses to refer to something in the world depends on how the speaker conceptualizes that stretch of the world. Word learning is not actually word-to-*world* mapping but word-to-*concept* mapping: The task of the child is not to map a word to a particular object or event in the world but to the concept under which the speaker has represented that object or event. Quine (1960/2013) illustrated this issue with the following thought experiment: Suppose that a stranger learning the language of a foreign country heard a native say "gavagai" while pointing to a running rabbit. What does "gavagai" refer to: the rabbit, the rabbit's ears, the act of running, a potentially delicious meal? The language learner must identify how the speaker conceptualized the scene in order to learn the meaning of this word.

Even in the realm of concrete nouns, the task of word-to-concept mapping can be quite difficult. When a speaker says "dog" in the presence of a furry domestic canine, how do children know that this word refers to the whole animal and not the dog's tail or whiskers? How do children know this label is not restricted to a particular breed of dog but could be extended to other dogs, although not to other types of pets or animals?

One theory proposes that children operate under learning biases that constrain the meanings they will hypothesize for a new noun. Markman (1994) posited three such biases: that a word will likely refer to an object kind rather than part of an object (the whole-object bias); that a word will likely refer to a basic-level category like "dog" rather than a subordinate category like "Dalmatian" or a superordinate category "animal" (the taxonomic bias); and that a word is not likely to label the same object that another known word already labels (the mutual exclusivity bias). All three of these biases have some degree of experimental support. Children generalize novel words presented as nouns to object kinds that share the same category, identifying the novel word as a label for the whole object rather than a part of the object (Balaban & Waxman, 1997; Markman & Hutchinson, 1984; Waxman & Markow, 1995; Woodward & Markman, 1998). They also prefer to generalize novel nouns to basic-level categories, such as "bird," rather than subordinate categories, such as "robin," or superordinate categories, such as "animal" (Hall & Waxman, 1993), and assume that a novel noun names an object for which they do not already have a word (Markman & Wachtel, 1988). Children's biases in word learning may therefore help them avoid the "gavagai" problem by restricting the range of concepts they think a new noun is likely to label.

Constraints like the whole-object, taxonomic, and mutual exclusivity biases might help children tackle the word-to-concept mapping problem for concrete nouns. But this problem appears vastly more difficult for verbs. These same biases do not immediately apply to verbs because verbs label events rather than objects (P. Bloom, 1994), and it is unclear whether analogous biases would apply to the way that verbs label events (Gleitman, 1990). Verbs were more difficult for the adults to identify than nouns in Gillette et al.'s (1999) word learning simulation, and several factors might conspire to create this difficulty. First, verbs are not necessarily uttered at the same time as the event they are describing but frequently are used to talk about past or future events instead. Beckwith, Tinker, and Bloom (1989) surveyed a corpus of maternal speech to children and found that the verb *open* was used 37.5% of the time to refer to something not in the present context. Second, the same event can be described by different verbs depending on the speaker's perspective: An event of a lion running after a gazelle could be described as the lion *chasing* the gazelle, or the gazelle *fleeing* the lion (Gleitman, 1990). A child attempting to learn whether a new verb means *chase* or *flee* would have little help from the context, because the contexts in which *chase* is used are identical to those in which *flee* is used. Finally, certain verbs describe events and states that cannot be observed at all: Attitude verbs, such as *think*, *want*, and *hope*, describe an individual's internal beliefs or desires, but these do not have observable physical correlates (Gleitman, 1990). Because observational learning appears insufficient in these cases, children must use different tools to overcome the challenges of verb learning.

Syntactic Bootstrapping

Children have another tool for acquiring verb meanings: the types of syntactic structures that verbs can occur in. If children know or can figure out the syntactic properties of a new verb, and they also know how those syntactic properties are related to verb meanings, then they might be able to infer aspects of the new verb's meaning. This strategy is called syntactic bootstrapping (Gleitman, 1990; Landau & Gleitman, 1985; Lasnik, 1989).

What relations between syntactic information and meaning might a child be able to exploit in verb learning? One type of syntactic information that is potentially easy to observe is the arguments in a sentence containing a verb. For example, a verb like *hit* can occur with a subject and an object in a sentence like *Sally hit her sister*. These arguments label participants in the event described by the sentence: The subject labels the person who did the hitting (the agent), and the object labels the person who got hit (the patient). Even if a child does not know the meaning of the word "hit," if that child is aware that subjects tend to name agents and objects tend to name patients, then she might infer that this sentence describes an event where Sally was the agent and Sally's sister was the patient.

Gertner, Fisher, and Eisengart (2006) found that children are able to use information about subjects and objects to infer which event a new verb labels. They played 2-year-olds a sentence with a novel verb, such as *The duck is gorping the bunny*, in the context of two different events: a scene with a duck pushing a bunny and a scene with the bunny pulling the duck's legs. The researchers used a method called preferential looking, which takes greater looking time toward one visual stimulus over another as evidence for how children understand a sentence. Children who heard *The duck is gorping the bunny* preferred to look at the scene where the duck was pushing the bunny: They interpreted *gorping* as pushing rather than leg-pulling. By identifying that *the duck* was the subject and *the bunny* was the object of *gorp*, these children were able to conclude that *gorp* named an event in which the duck was the agent and the bunny was the patient. Children were able to use the types of syntactic arguments that occurred with the novel verb to infer which event that verb labeled, given two options.

Not all verbs are able to occur in transitive sentences. Verbs such as *sleep*, *arrive*, and *fall* are intransitive: They occur in sentences with a subject only. That is, you can say *Doug fell* but not **Doug fell Andrew*. It appears that a verb's ability to occur in a transitive or intransitive sentence is related to the types of events it labels. Verbs like *hit*, *push*, and *bump* that occur in transitive sentences tend to label causative events with one participant acting on another. Verbs like *sleep*, *arrive*, and *fall* that occur in intransitive sentences tend to label noncausative events that have only one participant, such as the individual who is sleeping, arriving, or falling. Children might therefore exploit this correlation between clause type and event type in making an inference about the event a verb labels. In a preferential looking study, Naigles (1990) found that children who heard a transitive sentence, such as *The duck is gorping the bunny*, were more likely to look at a scene in which a duck pushed a bunny than at a scene in which a duck and a bunny wheeled their arms separately. By observing that the verb occurred in a transitive sentence, children inferred that it described a causative action: a pushing event rather than an arm-wheeling event. Naigles also found the reverse pattern for children who heard an intransitive sentence, such as *The duck and the bunny are gorping*: These children preferred the noncausative arm-wheeling event rather than the pushing event.

It appears that children are sensitive to the type of sentence containing a novel verb and will use that information to infer which event the verb labels, given two choices. One influential hypothesis proposed that children behave this way because they expect the number of arguments of a sentence to match one to one the number of participants in the event the sentence's verb describes (Fisher, Gertner, Scott, & Yuan, 2010; Gleitman, 1990; Naigles, 1990; Yuan, Fisher, & Snedeker, 2012). This hypothesis has been tested extensively, and children's tendency to infer that transitive verbs name causative 2-participant events has been replicated many times over, in children as young as 22 months old. (See, e.g., Fisher et al., 2010, for a review.) Lidz, Gleitman, and Gleitman (2003) corroborated this tendency in 3-year-old children learning Kannada, a Dravidian language spoken in south India, which has a verbal morpheme that signals when a verb is causative. This morpheme is a much more reliable cue to causative meaning than transitivity, because a verb can occur in a transitive sentence without necessarily having a causative meaning. Nonetheless, children acted out causative meanings for transitive verbs, even without the causative morpheme, and they acted out noncausative meanings for intransitive verbs, even with the causative morpheme. Even though these children speak a language that provides a more reliable morphological signal for causativity, they preferred to rely on transitivity when deciding whether to interpret a verb with a causative meaning.

Children therefore appear to use transitivity as evidence that verbs label causative 2-participant events. However, beyond Naigles's (1990) seminal experiment, later studies have not found consistent behavior with intransitive verbs: Children who hear intransitive sentences with novel verbs do not reliably prefer events with one participant, beyond what would be expected by chance (Arunachalam & Waxman, 2010; Noble, Rowland, & Pine, 2011; Yuan et al., 2012). These findings are puzzling under the hypothesis that children expect one-to-one matching between the arguments of a sentence and participants in the event the sentence describes. However, they are consistent with a weaker learning strategy: Perhaps children merely expect that each argument names a participant, but not necessarily vice versa (Williams, 2015). In this case, the sole argument of an intransitive sentence could name one of the participants in a 2-participant event, making both the 1- and 2-participant events possible referents for the novel verb. Further work is needed to determine the source of children's behavior with intransitive sentences, and what this reveals about the specific inference that children make when using clause type as evidence for verb meaning.

Children also may be able to draw inferences about verb meanings from other types of complements, or syntactic dependents that follow a verb. In the sentence *Sally hit her sister*, the complement of the verb *hit* is the noun phrase *her sister* (the object). But in the sentence *Jim thought that Gina liked him*, the complement of the verb *think* is a whole clause. If there are correlations between the types of complements a verb can take and the meanings a verb can express, then children might be able to use a verb's complement as evidence for its meaning.

Some initial evidence that complements might be useful sources of information comes from Landau and Gleitman (1985), who studied the language acquisition of English-speaking blind children. They found that these children acquire meanings for the verb *look* relative to their own haptic exploration rather than to sight: Blind children respond to the request to "Look up!" by reaching upward with their hands, whereas sighted children wearing a blindfold turn their heads upward. The meaning that the blind children have assigned to *look* is supported by the contexts in which this word is used: Blind children hear *look* in situations when a relevant object is nearby, and therefore haptic exploration is possible. Yet somehow these children manage to differentiate *look* from other verbs, such as *touch* and *hold*, which also are used when a relevant object is nearby but are interpreted as contact terms rather than perception terms. The researchers hypothesized that the syntactic distribution of *look* compared to *touch* and *hold* drives this difference in interpretation. *Look* takes different complements from *touch* or *hold*: You can say *look at that picture* but not **touch/hold at that picture,* and *look down* but not **touch/hold down*. Perhaps the blind children in this study used the differences in complements among these verbs to infer that they have different types of meanings, even though they heard them in the same physical contexts.

Further evidence that children use verb complements to draw inferences about verb meanings comes from the acquisition of attitude verbs, such as *think*, *want*, and *hope*. Recall that these verbs are particularly hard to learn because they name internal states of speakers' minds (Gleitman, 1990; Gleitman et al., 2005). A large body of literature also has argued that young children may have difficulty acquiring attitude verbs because they lack the mental state concepts that these verbs label; in particular, children fail in certain tasks to demonstrate the ability to represent others' beliefs (the so-called developing theory of mind, e.g., Astington & Gopnik, 1991; Flavell, Green, & Flavell, 1990; Gopnik & Wellman, 1994; Perner, 1991). However, more recent work finds that children's failure on these tests may be due to experimental and pragmatic factors rather than immature belief representations (e.g., Hansen, 2010; Z. He, Bolz, & Baillargeon, 2012; Helming, Strickland, & Jacob, 2014; Lewis, Hacquard, & Lidz, 2017; Onishi & Baillargeon, 2005; Rubio-Fernández & Geurts, 2013).

Yet even if children have the ability to represent speakers' mental states, learning which verbs label these mental states is not a trivial matter. These mental states do not have obvious physical correlates, and it is difficult to tell when mental states rather than actions are under discussion: If a speaker uses a new verb, how does a child know whether the verb labels what someone is feeling or what someone is doing? In the human simulation study by Gillette et al. (1999), adults were particularly bad at identifying attitude verbs from the visual contexts in which they were uttered; sometimes they could identify action verbs, such as *hit*, but they almost never identified attitude verbs, such as *think.* However, attitude verbs do have a reliable syntactic signal: their ability to take full clauses as complements. We can say *Jim thought that Gina liked him*, but not **Jim danced that Gina liked him*. Therefore, even though children may have difficulty identifying attitude verbs from the situational contexts in which they are used, children might be able to identify them through their syntactic distribution—specifically, by paying attention to which verbs take clausal complements (Fisher, Gleitman, & Gleitman, 1991; Gleitman et al., 2005).

Furthermore, differences in the clausal complements of attitude verbs might help children tell certain attitude verbs apart from each other. Attitude verbs fall into two major classes: Verbs like *think* and *know* convey meaning about speakers' beliefs, whereas verbs like *want* and *demand* convey meaning about speakers' desires. Cross-linguistically, these two classes of attitude verbs also differ in the properties of their clausal complements. In English, this difference is reflected in the tense (finiteness) of the complement. Desire verbs, such as *want*, tend to occur with nonfinite complements: We can say *I want John to be at home* but not **I want that John is at home.* By contrast, belief verbs, such as *think*, tend to occur with finite complements: We can say *I think that John is at home* but not **I think John to be at home*. The specific syntactic property that differentiates the complements of desire verbs from those of belief verbs varies across languages but seems to obey the following generalization: The complements of belief verbs look like declarative main clauses in each language, and the complements of desire verbs do not (Hacquard, 2014; White, Hacquard, & Lidz, 2017). If children are aware of this generalization, they might be able to use it to infer whether an attitude verb expresses a desire or a belief meaning.

A study by Harrigan, Hacquard, and Lidz (2016) found that 4-year-olds draw different inferences about the meaning of an attitude verb depending on whether they hear it with a finite or a nonfinite complement. The researchers tested the verb *hope*, which is special in its ability to take both types of complements: We can say *I hope to win the prize* or *I hope that I will win the prize*. This verb is also relatively rare in speech to children and is therefore less familiar than other attitude verbs, such as *think* or *want*. Without much prior verb knowledge to rely on, these preschoolers treated *hope* more like *think* when they heard it with a finite complement, and they treated it more like *want* when they heard it with a nonfinite complement. It seems that preschoolers can use the syntactic properties of an attitude verb's complement to infer whether the verb has a belief or a desire meaning.

With both action verbs and attitude verbs, children use their syntactic knowledge to overcome the challenges of word-to-concept mapping. Despite the difficulties of learning verb meanings by observation, children are able to identify aspects of their meanings by observing the syntactic structures that verbs occur in. Children use the arguments in a sentence to infer what type of event is labeled by a verb in that sentence, and they use more specific properties of a verb's complement to infer whether and how that verb labels hard-to-observe events, such as mental states. By identifying a verb's syntactic properties, and knowing something about how those syntactic properties map onto aspects of the verb's meaning, children can draw sophisticated inferences about the types of events that a verb can describe.

Summary

Children use a variety of tools to learn the lexicon of their language: both the grammatical categories of words and their meanings. Statistical sensitivities—tracking the distribution of words in the input and the extralinguistic contexts in which these words are used—can help learners group words into grammatical categories. Sensitivity to the contexts in which words are used can also help learners identify what some of these words refer to, at least for concrete nouns. But prior linguistic knowledge also intersects with statistical sensitivity to solve the missing pieces of the lexicon puzzle. Knowledge of the properties of grammatical categories helps children identify properties of a word's meaning that may not be identifiable just from the word's distribution in the input. Furthermore, knowledge of the syntactic structures in which a word can occur, and the ways in which syntactic structure maps onto meaning, helps children infer the range of concepts a new word can label, even if those concepts do not have observable physical correlates. Knowledge about the syntax of the child's language therefore plays an important role in word learning. Next we discuss how this syntactic knowledge is acquired.

SYNTAX

The system of rules in your language that allow you to combine words and morphemes into larger hierarchical structures is called syntax. Languages vary in some of their syntactic properties, such as the order of words and phrases. For example, children need to learn whether their language puts the subject before the verb and the object after the verb like English does, or whether the subject and object appear in a different order. However, languages do not seem to vary across other syntactic properties, such as organizing phrases into hierarchical structure. For example, in all languages, the subject of the sentence is structurally separate from the unit formed from the verb and the object (Baker, 2001). Furthermore, languages vary in which relations hold between specific elements of a sentence, but all languages encode these relations across hierarchical structures rather than linear strings of words (Chomsky, 1957, 1975). Children need to determine how their language behaves with respect to the syntactic properties that vary cross-linguistically but might take for granted the properties that do not vary, such as hierarchical structure and structure-dependent relations.

Clause Structure

Within a sentence, grammatical categories combine in specific ways into phrases, and these phrases combine into clauses with larger hierarchical structure. The order in which these units combine determines different hierarchical arrangements, with different meanings. For example, in English the noun *boy* can combine with a determiner *the* to form a noun phrase *the boy*. This noun phrase can be the object of a verb like *bite*, to produce a verb phrase: *bite* [*the boy*]. This verb phrase can combine with tense and agreement morphology as well as a subject noun phrase, such as *the cat*, to produce a full clause: *The cat* [*bites* [*the boy*]]. If these units were combined in a different order, we would get a different meaning: *The boy* [*bites* [*the cat*]].

In English, subjects generally precede verbs, which precede objects: English has "SVO" order. Knowing this word order allows you, as an adult speaker of English, to infer the structure of a sentence. In many sentences, if *the cat* comes before the verb *bite*, you can infer that *the cat* is the subject, whereas if *the boy* comes before *bite*, you can infer that *the boy* is the subject. However, this word order can vary across languages. Japanese tends to have SOV order, where the equivalent to *The cat bites the boy* would have an order like *The cat the boy bites*. Irish has VSO order: The Irish equivalent to this sentence would have an order like *Bites the cat the boy*. In order to assign a syntactic structure to sentences they hear, children must learn the property of their language that determines whether the subject and object precedes or follows the verb.

Children appear sensitive to their language's word order from a young age. As soon as children begin combining words, their utterances display the correct order of words in their language (L. Bloom, 1970; Brown, 1973). But even before they begin combining many words, children are able to infer properties of sentence structure from the order of words in a sentence. Hirsh-Pasek and Golinkoff (1996) played English-speaking 17-month-olds sentences in the context of two scenes: one showing Big Bird washing Cookie Monster and the other showing Cookie Monster washing Big Bird. Children who heard *Big Bird is washing Cookie Monster* looked more at the scene where Big Bird was doing the washing, and children who heard *Cookie Monster is washing Big Bird* looked more at the opposite scene. It appears that these children could identify the subject of the sentence based on its word order and inferred that the character labeled by the subject was the agent of the action. Recall that slightly older children could use this word order information to arrive at the correct interpretation of a novel verb: After hearing a sentence like *The duck is gorping the bunny*, they inferred that *gorp* must label an event where the duck was the agent rather than the bunny (Gertner et al., 2006). These children could identify where subjects and objects occur in sentences and use this information to make inferences about the sentence's meaning.

How do children develop this early understanding of word order? In the last section we discussed a strategy called semantic bootstrapping that could help children learn the grammatical categories of some words based on the types of meanings those words have. It also has been proposed that another form of semantic bootstrapping helps children infer the syntactic structure of sentences from the meanings of those sentences (Grimshaw, 1981; Pinker, 1984, 1989). Suppose a child hears a sentence like *The cat bites the boy* in the context of a scene where a cat bites a boy. If that child represents the scene as a biting event where the cat is the agent and the boy is the patient, and knows that the phrase *the cat* refers to the cat and *the boy* refers to the boy, she might be able to identify which phrases in the sentence are labeling the agent and which phrases are labeling the patient. If she furthermore knows that the agent of the event corresponds to the subject of a sentence and the patient corresponds to the object, she will be able to identify that *the cat* is the subject and *the boy* is the object. This information, combined with the knowledge that *bites* is a verb, will tell her that subjects come before verbs and objects come after verbs. She might then expect future English sentences to have SVO word order.

This semantic bootstrapping strategy relies on children being able to perceive scenes in the world under the right type of conceptual structure to align with the structure of sentences they are hearing. Prelinguistic infants appear to perceive events under conceptual structures that distinguish participant roles like "agent" from other participant roles like "patient." Children as young as 6 months represent agents as special participants in events, with intentions and goals (Csibra, Gergely, Bíró, Koos, & Brockbank, 1999; Leslie, 1995; Luo, Kaufman, & Baillargeon, 2009; Woodward, 1998). This agent role of an event therefore may be perceptually available for children to map onto a particular linguistic structure, such as the subject of a sentence. But in order to use this strategy, children do not just need to represent scenes with conceptual structure that can map onto sentence structure; they also need to represent those scenes in the same way as the speaker of the sentence did. Because a lion running after a gazelle could be described as either a "chasing" or a "fleeing" event (Gleitman, 1990), a child who hears *The gazelle fled the lion* to describe this scene would need to represent the scene as a fleeing event with the gazelle as the agent in order to infer that *the gazelle* is the subject of the sentence. If she instead represented the scene as a chasing event with the lion as the agent, she might mistakenly conclude that *the lion* is the subject in this sentence. Therefore, not all sentences will be equally informative about word order under this strategy.

This strategy also assumes that the child knows how agents and patients of events are represented linguistically in different structural positions in a clause: namely, that agents normally are represented as subjects and patients as objects. This pattern happens to be a very robust one cross-linguistically for active, transitive clauses (Baker, 1988; Dowty, 1991; Fillmore, 1968; Jackendoff, 1972). It is this "agents-are-subjects" expectation that allows our hypothetical learner to infer that *the cat* is the subject of *The cat bites the boy*, based on her knowledge that *the cat* labels the agent of the event being described. The 17-month-olds in Hirsh-Pasek and Golinkoff's (1996) study demonstrated knowledge of this generalization: They knew that when *Cookie Monster* appeared as the subject of the transitive sentence, the Cookie Monster character had to be the agent and not the patient of the action.

Further work has shown that children have such a strong expectation that the subject of a sentence will name the agent of the event that they have difficulty overriding this expectation in cases where the generalization does not hold. For example, children sometimes misinterpret passive sentences as active sentences, responding to a sentence like *The boy was bitten by the cat* by acting out or pointing to a situation where the boy bit the cat rather than a situation where the cat bit the boy (Bever, 1970; Turner & Rommetveit, 1967). It has been proposed that this behavior is due to children's strong expectation that the subject of the sentence names the agent of an event, along with difficulty detecting the cues that signal passive sentences or difficulty using those cues to revise initial interpretations (Bever, 1970; Huang, Zheng, Meng, & Snedeker, 2013; Li, Bates, & MacWhinney, 1993; Maratsos & Abramovitch, 1975; Stromswold, Eisenband, Norland, & Ratzan, 2002; Turner & Rommetveit, 1967); but see alternative interpretations in Borer and Wexler (1987, 1992), Brooks and Tomasello (1999), Demuth (1989), Gordon and Chafetz (1990), and F. N. Harris and Flora (1982). Children's expectation that agents are subjects may be such a useful principle in guiding the interpretation of basic clauses in their language that it sometimes leads them to make errors in interpreting nonbasic clauses.

Children therefore might be able to use principles like "agents-are-subjects" to identify the order of subjects and objects in their language, aligning their structured perception of events in the world with the structure of sentences describing those events. As long as children hear some clear cases where the sentences they hear align with their perception of the events being described, this semantic bootstrapping strategy might enable them to identify whether their language places the subject or object before or after the verb. But simply identifying that subjects, verbs, and objects occur in a particular order does not by itself tell the child that those units are arranged in a structural hierarchy—that verbs and objects form a unit and that subjects are structurally separate from that unit (e.g. Baker, 2001). How do children learn that sentences are built with this type of structure, with units built from smaller units? Because hierarchical structure is a feature of all human language, it is possible that children take this for granted: The language learning mechanism is constrained such that children acquiring any language will hypothesize only hierarchically structured syntactic representations (Chomsky, 1975).

Evidence for hierarchical structure in phrase representations has come from experimental work with children as young as 18 months old. Lidz, Waxman, and Freedman (2003) investigated whether these infants represented a noun phrase, such as *the yellow bottle*, as one big unit with no internal structure or whether *yellow bottle* forms a smaller unit inside the phrase: *the* [*yellow bottle*]. Adults have this nested representation, which is revealed in sentences like *I'll give Sarah this yellow bottle* and *I'll give you that one*. In this sentence, the word *one* does not refer just to a bottle—it refers to another *yellow* bottle. Because *one* can refer back to the string of words *yellow bottle*, those words must be a unit in the sentence. Lidz et al. showed 18-month-olds a picture of a yellow bottle and named it with a noun phrase that contained an adjective: *Look! A yellow bottle*. Then the infants saw a picture of another yellow bottle and a blue bottle, and heard either a sentence with the word *one* (*Do you see another one?*) or without (*What do you see now?*). Infants looked more at the yellow bottle than at the blue bottle, but only when they heard the word *one*. That is, they interpreted *one* to refer not to any bottle but specifically to another yellow bottle: They represented *yellow bottle* as a unit inside the phrase *a yellow bottle*. Even at very early stages of syntactic development, children's syntactic representations contain hierarchical structure.

Children can identify where subjects and objects occur in a sentence in their language and represent sentences with hierarchical structure even before they are producing many full sentences of their own. However, once children do begin producing sentences, two characteristics of their speech have led researchers to question the completeness of their sentence representations. One of these phenomena is the so-called root infinitive stage of early child speech, in which young children use the infinitive form of a verb instead of the tensed form. Because there are links between the morphological form of a verb and its position in a clause, a sizable literature has investigated whether children's root infinitive productions reflect immature knowledge about where verbs and other functional elements occur in the hierarchical structure of a clause (Bar-Shalom & Snyder, 1997; Guasti, 2002; Guilfoyle & Noonan, 1988; Haegeman, 1995; T. Harris & Wexler, 1996; Legate & Yang, 2007; Phillips, 1995; Platzack, 1990; Poeppel & Wexler, 1993; Radford, 1990; Rizzi, 1993, 1994; Schaeffer & Ben Shalom, 2004; Weverink, 1989; Wexler, 1994, 1998). The cause of this phenomenon remains a puzzle, but children generally pass through the root infinitive stage before they are 3 years old.

A second phenomenon is young children's omission of overt subjects in languages that require them, such as English. A large body of literature has investigated whether these early subject omissions reflect immature knowledge about the property of English main clauses that makes overt subjects obligatory or whether they reflect the interaction of other cognitive and linguistic factors, such as immature working memory, pragmatics, and prosody (Allen, 2000; P. Bloom, 1990; Gerken, 1991, 1994; Guasti, 2002; Hyams, 1986, 1992, 2011; Hyams & Wexler, 1993; Kim, 2000; Rizzi, 1993, 1994; Serratrice, 2005; Valian, 1991; Valian & Aubry, 2005; Valian & Eisenberg, 1996; Valian, Hoeffner, & Aubry, 1996; Wang, Lillo-Martin, Best, & Levitt, 1992). The source of children's early subject omissions is still under debate, but by the age of 3 children produce overt subjects consistently in languages that require them.

To summarize our discussion so far, children's acquisition of the clause structure of their language is informed by domain-specific constraints on their linguistic representations, interacting with properties of their perceptual system. Alignment between children's conceptual representations of events around them and the structure of at least some sentences that describe those events may help children identify which phrases label agents or patients. If children then expect that agents are subjects of transitive clauses, they can identify how their language orders subjects and objects with respect to the verb in a sentence. Although the completeness of children's early sentence representations has been debated, the structural hierarchy of subjects, verbs, and objects within a sentence, as well as the structural hierarchy of words within a single phrase of a sentence, may be something children take for granted. Hierarchical structure is common to all the world's languages, and children at the earliest stages of syntactic development appear to have hierarchically structured phrase representations. The requirement that linguistic expressions contain hierarchical structure therefore may be an innate constraint imposed by children's language learning mechanism.

Syntactic Dependencies

Syntactic dependencies are relations between elements in a clause or across clauses, determined by the syntactic properties of those elements and the structures they occur in.

Here we consider how children acquire two types of dependencies. The first type occurs in the sentence *Jane is playing the piano*. In this sentence there is a dependency between the auxiliary verb *is* and the *-ing* form of the verb, which work together to tell you that the sentence is in the present progressive, so Jane's playing is ongoing. This type of relation can hold across intervening material, as in the sentences *Jane is softly playing the piano* and *Jane is softly and beautifully playing the piano.* Because this dependency holds between two morphemes in a certain syntactic relation, it is a type of morphosyntactic dependency. A second type of dependency occurs in questions like *Which sonata is Jane playing tonight?* Here there is a dependency between the "wh-phrase" *which sonata* and the verb *playing*: We understand this question as asking about the missing object of that verb. We also find this type of relation in a relative clause like *I love the sonata that Jane is playing in the concert*, and this relation can hold across a lot of intervening material: *I love the sonata that Tony thought the program said that Jane is playing in the concert.* Because the object of the verb appears to have moved to a different position in these sentences, these dependencies are called movement dependencies. They also frequently are called filler-gap dependencies because the moved element is a filler that becomes associated with a gap later on in the sentence.

An important feature of these relations is the fact that they are defined over the hierarchical structure of elements in a sentence (Chomsky, 1975). In other words, the relations that elements of a sentence can enter into depend on their structural positions with respect to each other. For instance, the dependency between *is* and *-ing* does not hold between *is* and any sequence of sounds pronounced "ing" that occurs after it: We do not get this dependency in sentences like *Jake is a singer* or *The probability is vanishingly small*. We get this dependency only between *is* and an *-ing* morpheme that occurs on the main verb in the sentence. This dependency is defined over a particular structural relation between *is* and *-ing*, not the linear order of these two sounds. And it is not the case that any string of words can enter into a movement dependency in a wh-question or relative clause; only strings that are units within the hierarchical structure of the sentence can move. In the sentence *Jane is playing which sonata in the concert?*, the string of words *which sonata* is a unit, so it can move to the front of the sentence, creating a movement dependency: *Which sonata is Jane playing in the concert?* However, *which sonata in* is not a unit, so those words cannot move together: We cannot say **Which sonata in is Jane playing the concert?* Movement dependencies are constrained by the hierarchical structure of the sentence and can hold only between structural units in the sentence.

Morphosyntactic Dependencies

Children's statistical sensitivities and extralinguistic cognition interact to help them identify morphosyntactic dependencies in their language. Experimental work with very young children has found that they can track the statistical signature of dependencies like the *is-ing* relation, but this ability is mediated by their memory resources. Santelmann and Jusczyk (1998) played 18-month-olds sentences with the sequence *is* Verb-*ing*, a real English dependency, as well as sentences containing the sequence *can* Verb-*ing*, which is not an English dependency. For example, some children heard sentences like *Everybody is baking bread*, and other children heard sentences like **Everybody can baking bread.* Eighteen-month-olds preferred to listen to sentences with the *is* Verb-*ing* sequence over sentences with the *can* Verb-*ing* sequence, indicating that they knew that *is* and *-ing* signal a real morphosyntactic dependency in English. These children also preferred sentences with *is* Verb-*ing* when a 2-syllable adverb came between *is* and the verb, but not when a longer adverb intervened: They still were able to detect this dependency in sentences like *Everybody is often baking bread* but not in *Everybody is effectively baking bread*. It appears that these infants' limited memory resources interfered with their ability to detect the signal of this morphosyntactic dependency. That is, children needed to be able to hold enough linguistic material in memory in order to detect the co-occurrence of *is* with *-ing*, and longer intervening adverbs taxed their limited memory resources enough to prevent them from detecting this dependency.

Santelmann and Jusczyk's (1998) results indicate that English-speaking children are aware of the morphosyntactic dependency between *is* and *-ing* by the age of 18 months, although their memory resources are not always sufficient to detect this dependency in their input. What allows children to become aware of this dependency? Results from artificial language learning studies suggest that children can track co-occurrence patterns in their input to learn nonadjacent dependencies, such as the one between *is* and *-ing* in English (Gómez, 2002; Gómez & Maye, 2005). Recall that in our discussion of word segmentation, young children could use statistics to track the probability that certain nonsense syllables would occur next to each other (e.g., Saffran et al., 1996). Now the question is whether children can track the probability that certain strings will occur together across intervening material—for example, that *is* will co-occur with *-ing* with different verbs in between. Gómez and Maye (2005) tested 15-month-olds' abilities to detect these types of nonadjacent dependencies in an artificial language. These children heard "sentences" like *pel-vamey-rud*, *pel-wadim-rud*, and *pel-tapsu-rud*, in which a dependency between the nonwords *pel* and *rud* obtained across a variety of intervening nonwords. After training, these infants were able to recognize this *pel*-X-*rud* dependency in new "sentences" that contained it, as long as their training contained enough variety in the nonwords that came between *pel* and *rud*. This finding suggests that children as young as 15 months old are able to detect the statistical signature of nonadjacent dependencies, provided they hear enough variety in the intervening material.

Because morphosyntactic dependencies like the one between *is* and *-ing* in English are defined over hierarchical structures in a sentence rather than over the linear order of words, these relations can hold across large amounts of intervening material. Children's ability to detect the statistical signatures of nonadjacent dependencies is therefore crucial for learning these morphosyntactic dependencies in their language. But these statistical sensitivities interact with their extralinguistic cognition: Children need sufficient memory resources to recognize these dependencies over longer distances and may be unable to keep both parts of the dependency in memory if the amount of linguistic material between them grows too large. Children's ability to detect morphosyntactic dependencies in their language develops as their memory resources mature.

Movement Dependencies

Learning movement dependencies similarly involves interaction among children's statistical sensitivities, extralinguistic cognition, and domain-specific biases. We have seen that movement dependencies can hold only between structural units in a sentence. Because this structure dependence is a universal property of human language, it is something that children might take for granted: It might be an intrinsic constraint imposed by their language learning mechanism (Chomsky, 1975). This constraint would provide useful guidance for learning movement dependencies in their language: Once children can identify the hierarchical structure of a sentence, they will know that only units within that structure can move, and therefore they will know which instances of movement are possible and impossible.

Takahashi and Lidz (2008) and Takahashi (2009) used an artificial language learning paradigm to test children's knowledge of structure dependence. Following a method developed by Thompson and Newport (2007), they constructed artificial grammars in which some sequences of nonsense word categories could be optional, repeated, or substituted by other categories, which affected the probabilities of certain word categories occurring after others. After being trained on this artificial language, adults and 18-month-olds were tested on sentences that contained movement. Adults accepted sentences when one of the optional, repeated, or substituted category sequences was moved: They used the differences in transitional probabilities to group these sequences into units and recognized that those units could move. Eighteen-month-olds likewise accepted sentences with moved units and showed surprise when they heard sentences with moved sequences that were not units. In other words, these infants knew that only strings of words that form a unit within a structural hierarchy could take part in movement relations, even though they had never heard movement before in this task. Once they were able to identify the hierarchical structure of these sentences, they were able to identify possible and impossible instances of movement in this artificial language. Their knowledge of structure dependence allowed these learners to draw conclusions about syntactic relations beyond what they were exposed to in their input.

But knowing which elements of a sentence can and cannot move is only one step in learning movement dependencies. Children also need to be able to identify when this movement happens in sentences they hear. When adults hear a filler (a moved word) in a sentence, they quickly identify that the sentence contains a movement dependency and predict gaps where that filler could be interpreted (Crain & Fodor, 1985; Frazier & Clifton, 1989; Frazier & d'Arcais, 1989). Children are able to parse certain wh-questions in this predictive manner by the age of 5. Omaki et al. (2014) asked 5-year-olds questions like *Where did Lizzie tell someone that she was gonna catch butterflies?* and found that children interpreted the wh-word *where* as describing the location of the first verb that it could be associated with. English-speaking children interpreted this sentence as a question about the location of telling, and Japanese-speaking children interpreted the Japanese analog as a question about the location of catching, because the verb for "catch" comes before the verb for "tell" in Japanese word order. These children did not wait to hear the full structure of the sentence before resolving the movement dependency: They predicted that the wh-word could be interpreted with the first verb they encountered. In order to do this, children needed to detect cues in the sentence that told them a filler was present, predict upcoming structure, and keep the filler in memory while hearing the rest of the sentence, so they could access it and integrate it into their sentence representation as soon as possible. Children's developing extralinguistic cognition, in addition to their developing linguistic knowledge, might mediate their ability both to detect cues to movement dependencies and to resolve these dependencies accurately.

Some studies have found suggestive evidence that English-learning children develop the ability to detect movement dependencies in English sentences between the ages of 15 and 20 months (Gagliardi, Mease, & Lidz, 2016; Seidl, Hollich, & Jusczyk, 2003). Gagliardi et al. (2016) used a preferential looking method to test comprehension of wh-questions like *Which dog did the cat bump?* and relative clauses like *Find the dog that/who the cat bumped*. They found an interesting U-shaped learning pattern. Fifteen-month-olds appeared to arrive at the correct interpretation for both types of sentences: They looked more at a dog that got bumped than at a dog that was the agent of bumping. But 20-month-olds appeared to comprehend only wh-questions and relative clauses with *who*, not relative clauses with *that*. Twenty-month-olds' surprising failure with certain relative clauses might demonstrate the development of syntactic knowledge: They have learned to represent the full movement dependencies in these sentences but have difficulty detecting when relative clauses with *that* contain these dependencies. The word *that* is ambiguous in English—it occurs in many contexts other than in relative clauses—so words like *who* or *which* are much clearer cues to movement dependencies. Fifteen-month-olds, in contrast, might be arriving at the right answer through a heuristic that does not require them to parse the full movement dependency, thereby avoiding these difficulties with relative clauses.

Relative clauses therefore might pose challenges to the parsing mechanisms of early learners. However, children's difficulty in comprehending relative clauses throughout development has led many researchers to question whether children's linguistic representations of these dependencies are at fault. Children produce relative clauses as young as 2 years of age (Corrêa, 1995; Guasti, Dubugnon, Hasan-Shlonsky, & Schneitter, 1996; Labelle, 1990; McKee, McDaniel, & Snedeker, 1998), but even through their preschool years, children have difficulty comprehending some types of relative clauses when asked to act them out or point to a matching picture (de Villiers, Flusberg, Hakuta, & Cohen, 1979; Goodluck & Tavakolian, 1982; Hamburger & Crain, 1982; Sheldon, 1974; Tavakolian, 1981). Preschoolers have particular difficulty with relative clauses in which the filler is interpreted as the object rather than the subject of the verb. This is the difference between relative clauses like *the dog that the cat bumped* and *the dog that bumped the cat*: In the first, *the dog* is interpreted as the object of *bump*, and in the second, it is interpreted as the subject. Some researchers have attributed children's difficulty with object relative clauses to immature representations of these sentences (Labelle, 1990; Tavakolian, 1981), inability to represent certain object relatives that children have not frequently heard before (Arnon, 2009; Brandt, Kidd, Lieven, & Tomasello, 2009; Kidd & Bavin, 2002; Kidd, Brandt, Lieven, & Tomasello, 2007), or non-adult-like constraints on when this type of movement dependency can occur (Adani, Forgiarini, Guasti, & Van der Lely, 2014; Adani, Van der Lely, Forgiarini, & Guasti, 2010; Belletti, Friedmann, Brunato, & Rizzi, 2012; Friedmann, Belletti, & Rizzi, 2009). However, the sentence processing literature has found that adults also have difficulty with object relatives, reading them more slowly than subject relatives, which has been attributed to constraints on how memory is accessed in resolving these dependencies. (See Wagers & Phillips, 2014, for a review.) It therefore is possible that preschoolers' difficulties in interpreting these sentences might stem from the same type of memory constraints (e.g., Arosio, Yatsushiro, Forgiarini, & Guasti, 2012; Haendler, Kliegl, & Adani, 2015).

In summary, in order to learn syntactic dependencies, children must both detect these dependencies in their input and arrive at a structural representation for them. Children's extralinguistic cognition interacts with their domain-specific linguistic biases during both of these steps. Statistical sensitivities help children detect which morphemes are involved in a morphosyntactic dependency and which types of words signal that a movement dependency is present. Memory resources also contribute to this process, because children must be able to maintain and access linguistic information in memory in order to recognize and resolve dependencies that hold across intervening material. But domain-specific biases also play an important role. If children take for granted that dependencies are defined over hierarchical structure, it may make it easier to learn them: Once children have identified the structure of a clause, they will have information about what types of dependencies can hold between elements in that structure.

Summary

Children use their extralinguistic perceptual and memory systems, statistical sensitivities, and domain-specific knowledge about the nature of linguistic representations to learn the syntax of their language. Children might be able to infer which phrases of their language are subjects and objects by observing whether those phrases label agents or patients in their representations of events in the world. Children's statistical sensitivities and memory resources help them detect and resolve syntactic dependencies between elements of a sentence. But children may not need to learn that elements in a sentence are arranged in a structural hierarchy or that syntactic dependencies operate over this hierarchical structure: These are syntactic properties common to all of the world's languages, so children might take them for granted. The nature of the language learning mechanism constrains children's early linguistic representations to be hierarchically structured and constrains children to posit syntactic dependencies only between units in this structure. Knowledge about the syntactic properties of human language in general therefore allows children to draw inferences about the syntactic structure of their own language.

SEMANTICS

Semantics is the study of how linguistic expressions convey meaning. The meaning of a sentence is more than just a sum of the meanings of the words but depends on sentence structure as well. We have seen that sentences like *The dog chased the cat* and *The cat chased the* dog convey different meanings despite having all the same words. These sentences have different meanings because the role that each of the noun phrases plays is different in each sentence.

Sentence structure contributes to many aspects of sentence meaning, not just role assignment. For example, the way that pronouns are interpreted depends on their syntactic context. Pronouns make a contribution to sentence meaning that is underspecified. Assigning an interpretation to the pronoun (and hence to the sentence) depends on the context of use. In the sentence *Allison thinks that she will get the job*, the pronoun can be interpreted as referring either to Allison or to some other salient individual in the context.

When a pronoun gets its interpretation based on the interpretation of some other phrase, the relation between the two expressions is subject to syntactic conditions. For example, the pronouns *she* or *her* may get their reference from (corefer with) *Belinda* in sentences like *When she was in the interview, Belinda spilled some water* and *Belinda said that my brother interviewed her*. But the pronouns all must refer to someone other than Belinda in sentences like *She was in the interview when Belinda spilled some water* and *Belinda interviewed her.* Thus, while we can characterize pronouns as those expressions whose reference can be determined by other parts of the sentence, the conditions under which such referential dependencies hold are constrained by syntax in ways that we discuss in the following section "Interpreting Pronouns."

Other kinds of semantic relations between words and phrases also are dependent on properties of sentence structure. For example, the sentence *Every student didn't complain about his grades* is ambiguous. This sentence can express the idea that no students complained. It also can express the weaker idea that some students complained and others did not. This ambiguity arises because of the relative scope of negation and the universal quantifier *every*. Scope is the domain in which a quantifier or other operator can influence how other expressions are interpreted. In this sentence, *every* can be interpreted either outside or inside the scope of negation. If the sentence means "every student is such that he didn't complain about his grades," we get the stronger reading, whereas if it means "not every student complained about his grades," we get the weaker meaning.

In the remainder of this section, we consider the acquisition of constraints on pronoun interpretation and on quantifier scope. These issues have been a focus of research in language acquisition because they reveal the highly abstract nature of the rules governing the interpretation of sentences and thus highlight the potential disconnect between the nature of experience and acquired grammatical knowledge.

Interpreting Pronouns

Pronouns can fix their reference through some other noun phrase, but there are constraints on the kinds of sentences in which this can happen. These constraints are based on two factors: structural hierarchy and structural locality.

The role of hierarchy can be seen in the contrast between *When she was in the interview, Belinda spilled some water* and *She was in the interview when Belinda spilled some water*. In each of these sentences, the pronoun precedes *Belinda* in the linear order of words, but in the second sentence, the pronoun is "higher" in the structural hierarchy. The notion of height in linguistic structures is expressed though a relation called c-command (Reinhart, 1981). One expression c-commands another if the smallest unit containing the first also contains the second. In the first sentence provided, the pronoun does not c-command *Belinda*, but in the second sentence, it does. In addition, one expression binds a second expression if it c-commands the second expression and corefers with that expression (Chomsky, 1981). But we cannot interpret the second sentence above with the pronoun coreferring with *Belinda*: It has to refer to someone else. In other words, the pronoun cannot bind *Belinda.* The relevant constraint on pronoun interpretation, known as Principle C, is thus that a pronoun cannot bind its antecedent (Lasnik, 1976); stated slightly differently, a referring expression like *Belinda* cannot be bound (Chomsky, 1981).

The structural notion of locality, when combined with c-command, explains the contrast between *Belinda said that my brother* *interviewed her* and *Belinda interviewed her*. In both sentences, *Belinda* c-commands the pronoun, but only the first allows coreference. This is because of the locality condition, known as Principle B (Chomsky, 1981), requiring that a pronoun not be bound in the smallest clause containing it. In the first sentence, the pronoun and *Belinda* are in different clauses, so *Belinda* can bind the pronoun and the two expressions can corefer. But in the second sentence, the two expressions are in the same clause, so *Belinda* cannot bind the pronoun: The coreferential interpretation is ungrammatical. Instead, the pronoun must refer to someone else.

Early work on the acquisition of Principle B found that children as old as 5 were sensitive to c-command but not to locality (Chien & Wexler, 1990), and hence allowed coreference in sentences like *Belinda interviewed her.* (See also Grodzinsky & Reinhart, 1993; Thornton & Wexler, 1999, among others.) However, Conroy, Takahashi, Lidz, and Phillips (2009) found that children do respect the locality portion of Principle B, and they argued that earlier results derived from methodological artifacts and biases coming from online sentence processing.

Principle C has played a very prominent role in arguments concerning the origins of grammatical knowledge. Because children are exposed only to sentence-meaning pairs that are grammatical, it is a puzzle how they acquire constraints like Principle C, which bar certain sentences from expressing otherwise sensible interpretations. How can one acquire rules about the interpretations that sentences cannot have?

Crain and McKee (1985) examined English-learning preschoolers' knowledge of Principle C, asking whether children know that a pronoun can precede its antecedent but cannot c-command it. The experimenters used a truth-value judgment task, in which participants observe a story acted out by the experimenter with toys and props. At the end of the story, a puppet makes a statement about the story. The participant's task is to tell the puppet whether he was right or wrong. Crain and McKee presented children with sentences like *While he was dancing, the Ninja Turtle ate pizza* and *He ate pizza while the Ninja Turtle was dancing* following stories with two crucial features. First, the Ninja Turtle ate pizza while dancing. This makes the interpretation in which the pronoun (*he*) and the referring expression (*the Ninja Turtle*) are coreferential true. Second, there was an additional salient character who did not eat pizza while the Ninja Turtle danced. This aspect of the story makes the interpretation in which the pronoun refers to a character not named in the test sentence false. Thus, if children allow coreference in these sentences, they should accept them as true, but if children disallow coreference, then they should reject them as false. The reasoning behind this manipulation is as follows. If children reject the coreference interpretation, then they must search for an antecedent for the pronoun outside of the sentence. Doing so, however, makes the sentence false.

Crain and McKee found that children as young as 3 years old accepted sentences like *While he was dancing, the Ninja Turtle ate pizza* in contexts that made the coreferential interpretation true but overwhelmingly rejected sentences like *He ate pizza while the Ninja Turtle was dancing* in identical contexts. The fact that they treated the two sentence types differently, rejecting coreference only in those sentences that violate Principle C, indicates that by 3 years of age, English-learning children respect Principle C.

The observation that Principle C constrains children's interpretations raises the question of the origin of this constraint. The fact that children as young as 3 years of age behave at adult-like levels in rejecting sentences that violate Principle C often is taken as strong evidence for the role of c-command in children's representations and hence for the role of hierarchical structure in shaping children's interpretations. (See Kazanina & Phillips, 2001, for supporting evidence from Russian.)

This view may be further bolstered by work demonstrating that children as young as 30 months display knowledge of Principle C. Lukyanenko, Conroy, and Lidz (2014) conducted a preferential looking experiment in which infants saw two videos side by side. In one video, a girl (Katie) was patting herself on the head. In the other video, a second girl patted Katie on the head. Infants were then asked to find the image in which "She is patting Katie" or the one in which "She is patting herself." Infants in the former condition looked more at the video in which Katie was getting patted by someone else, whereas those in the latter condition looked more at the video in which Katie was patting herself.

To determine whether children's interpretations were driven by Principle C, as opposed to an alternative nonstructural heuristic, Sutton, Fetters, and Lidz (2012) and Sutton (2015) tested children in a preferential looking task like Lukyanenko et al. (2014) and also in a task measuring sensitivity to hierarchical structure. Children saw three objects: a big red train, a medium-size yellow train, and a small yellow train. They then were asked to find "the big yellow train." Correct interpretation requires restricting the adjective *big* to apply to the phrase *yellow train*. Sutton et al. measured the speed with which the children looked to the correct object and used that to predict the speed with which they arrived at the correct interpretation of the Principle C sentences. They found that these structural processing measures were significantly correlated, although measures of lexical processing speed and vocabulary size were not predictive of Principle C performance. Together these findings suggest that the computation of hierarchical structure is a critical component of children's understanding of sentences, which are subject to Principle C from the earliest stages of syntactic development.

Quantification and Scope

Some sentences with quantifiers permit readings that do not follow directly from simple mapping of surface form to semantic interpretation (Büring, 1997; Horn, 1989; Jackendoff, 1972, among others). Consider the sentence *Every horse didn't jump over the fence*. This sentence is scopally ambiguous. On the interpretation that "every horse is such that it didn't jump over the fence," the sentence means that none of the horses jumped over the fence. Here *every* takes scope over negation. We call this an isomorphic interpretation because the scope relation between *every* and negation coincides with their surface positions. Another possible interpretation is that "not every horse jumped over the fence," which means that some horses jumped and some did not. In this case, negation takes scope over *every*. We call this a nonisomorphic interpretation because here negation takes scope over the whole sentence (i.e., in a position different from the one it occupies in surface syntax).

Musolino, Crain, and Thornton (2000) tested children's comprehension of quantificationally ambiguous sentences. They found that while adults can easily access the nonisomorphic interpretations of such sentences, 4-year-olds systematically assign such sentences an isomorphic interpretation only. This was true also for sentences like *The Smurf didn't buy every orange*, in which the isomorphic reading is the opposite from that of *Every horse didn't jump over the fence*. In the first sentence, 4-year-olds interpret negation as scoping over *every*, taking the sentence to mean "it is not the case that the Smurf bought every orange." In the second sentence, they interpret *every* as scoping over negation, taking the sentence to mean "every horse is such that it didn't jump over the fence." The authors take the finding that children systematically assign these sentences isomorphic interpretations to conclude that young children, unlike adults, systematically interpret negation and quantifiers on the basis of their position in overt syntax.

Musolino et al.'s (2000) findings, however, do not tell us the nature of the representations underlying children's resistance to nonisomorphic interpretations. One possibility is that children's overly isomorphic interpretations reflect the linear arrangement between quantifiers and negation. Alternatively, children's interpretations may be constrained by the surface c-command relations holding between these elements. These alternatives arise because c-command and linear order are systematically confounded in the materials used by Musolino et al.

In order to tease these possibilities apart, Lidz and Musolino (2002) compared English-speaking children's scope interpretations with those of Kannada-speaking children. The canonical word order in Kannada is subject-object-verb (SOV), and Kannada displays the same kind of scope ambiguities as English with respect to negation and quantifiers (Lidz, 2006). These properties are illustrated in sentences like *Naanu eraDu pustaka ood-al-illa* ("I didn't read two books"), which has the word order "I two books read not." This can mean "it is not the case that I read two books," where negation takes scope over the numeral, or "there are two books that I did not read," where the numeral takes scope over negation.

The crucial difference between Kannada and English is that in Kannada, linear order and c-command are not confounded. In both languages, negation c-commands the direct object in the structure of the sentence. However, the linear order of the words is different in the two languages: Negation precedes the object in English but follows the object in Kannada. Lidz and Musolino (2002) found that children interpret sentences like *The Smurf didn't catch two guys* with negation taking scope over the numeral, independent of the language being acquired. This finding illustrates that children's scope assignment preferences reflect the hierarchical relation of c-command and not merely the linear order of words.

Subsequent work on children's scope assignment reveals that their limitations likely derive from the pressures of online sentence understanding. First, Musolino and Lidz (2006) showed that children can access nonisomorphic interpretations when they are heavily supported by the discourse. These authors found a significant increase in nonisomorphic interpretations in contrastive contexts like *Every horse jumped over the log but every horse didn't jump over the fence.* Viau, Lidz, and Musolino (2010) went on to show that experience with contrastive contexts make children more readily accept nonisomorphic interpretations even in noncontrastive contexts. These results suggest that children's difficulties have more to do with deploying their knowledge in real time than with acquiring that knowledge in the first place.

Origins of Quantifier Meanings

By the age of 4, children have acquired the complex mapping between syntactic hierarchy and semantic interpretation in language. But how are quantificational terms acquired to begin with? Here we consider the cognitive and linguistic resources that contribute to quantifier acquisition.

Humans have multiple ways of representing information that is relevant for quantification. First, we have an ability to approximate the number of items in a scene through the approximate number system (ANS; Dehaene, 2009; Feigenson, Dehaene, & Spelke, 2004; Whalen, Gallistel, & Gelman, 1999). The ANS is a system that provides nonexact representations of cardinality, is present in infancy (Izard, Sann, Spelke, & Streri, 2009; Xu & Spelke, 2000), and increases in acuity throughout development (Halberda & Feigenson, 2008; Halberda, Ly, Wilmer, Naiman, & Germine, 2012). Between the ages of 3 and 4, children also acquire a system of precise cardinality, whereby they can represent the number of items in a scene exactly and refer to that quantity with number words (Carey, 2009; Gelman & Gallistel, 1978; Wynn, 1992). Finally, infants also can represent sets (Feigenson & Carey, 2003) and can keep track of multiple sets, allowing them to increase the number of individuals they can track in memory (Feigenson & Halberda, 2004, 2008).

Halberda, Taing, and Lidz (2008) asked whether children required knowledge of precise number in order to acquire the meaning of *most*, whose meaning depends on numerosity. They found that acquisition of precise number concepts is not a prerequisite for acquiring *most*. Many children acquire *most* prior to learning precise cardinality. Odic, Halberda, Pietroski, and Lidz (n.d.) went on to show that many children who have just acquired precise cardinality concepts and who have just counted the items in an array nonetheless will use the ANS to answer questions like *Are most of the animals giraffes or lizards?* This finding suggests that early acquisition of *most* is grounded in the ANS as a way of measuring cardinality. Odic et al. (2013) extended these results to *more*, showing that children acquire *more* at around 3.5 years, prior to many children's acquisition of precise cardinality.

Properties of children's cognitive systems for representing number therefore affect their early interpretations of quantifiers. But their linguistic knowledge may help them identify which words are quantifiers and therefore should receive a quantity-based meaning. Recall that the 4-year-olds in Wellwood et al. (2016) assigned a novel word like *gleebest* a quantity-based interpretation when it occurred in the syntactic position of a determiner or quantifier (e.g., *gleebest of the cows*) but assigned it a quality-based interpretation when it occurred in the syntactic position of an adjective (*the gleebest cows*). Children therefore can use their knowledge of the distribution of quantificational elements to infer that words that distribute like quantifiers must express quantificational meanings.

Preschoolers also appear to be sensitive to a subtler property of quantifier meanings that holds true cross-linguistically. Think about the sentence *Every girl is on the beach*. In order to assess whether this sentence is true, all you have to do is consider the set of girls in the discourse and see whether they are all on the beach. You do not have to consider boys or anything else in the discourse that is not a girl. This is due to a property of *every* called conservativity, and it is a property shared by all quantifiers in human language (Barwise & Cooper, 1981; Higginbotham & May, 1981; Keenan & Stavi, 1986).

Hunter and Lidz (2013) investigated 4- and 5-year-olds' knowledge of conservativity by seeing whether children could learn a novel quantifier that did not have this property. Children were trained to select cards that corresponded to the meaning of a novel quantifier (*gleeb*). In one case, the intended meaning of *gleeb* was "not all": Children were shown that *Gleeb girls are on the beach* only matched cards where not all girls were on the beach. In this case, *gleeb* is conservative because only the girls need to be considered in order to verify the sentence. In the other case, *gleeb* meant the mirror image of "not all": *Gleeb girls are on the beach* only matched cards where not all people on the beach were girls. This version of *gleeb* is not conservative because all of the beach-goers, not just the girls, need to be considered in order to verify the sentence. After training, children showed evidence of learning the conservative *gleeb* but failed to learn the nonconservative *gleeb*. This finding suggests that preschoolers know that words presented in quantifier contexts must have conservative meanings as a consequence of being quantifiers.

Thus, children's acquisition of quantifier meanings is influenced by both linguistic and extralinguistic factors. Properties of children's developing cognitive systems for representing number affect how they interpret words whose meaning depends on numerosity, but prior linguistic experience with the distribution of quantifiers helps children infer which words have number-based meanings to begin with. Further domain-specific linguistic constraints restrict the types of meanings for quantifiers that children will consider.

Summary

When we examine children's acquisition of the constraints on sentence interpretation, we see strikingly specific and early knowledge of the ways sentence structures can map to possible sentence meanings. The interpretations that children assign to pronouns, and the interpretations that children avoid, reveal their knowledge of the cross-linguistic constraints on when pronouns can corefer with other noun phrases in certain structural configurations. Children's interpretations of quantifiers reflect sensitivity to the structural positions of quantifiers within a sentence as well as the possible quantifier meanings that human languages allow. This linguistic knowledge interacts with children's extralinguistic cognitive systems, which influence their ability to process complex sentence structures online and represent the number concepts that quantifiers express. But because children's early knowledge of the interpretations that pronouns and quantifiers cannot have cross-linguistically would be extremely difficult to acquire by observing the interpretations that are possible in their language, this knowledge likely stems from constraints inherent to their linguistic system. Children's early semantic knowledge is therefore particularly revealing about the rich structure of the mechanism that guides their language learning process.

CONCLUSION

Within the first 6 years of their lives, children develop the ability to speak with and understand those in their community by acquiring a shared cognitive system—the grammar of their language—that links speech sounds with meanings. Just as our visual faculty is exposed to light and interprets that signal to infer the structure of the object that the light is reflecting off of, our language faculty, when exposed to speech sounds, interprets those signals to infer the structure and meaning of the sentence underlying those sounds. In order to acquire the ability to map sounds to meanings, the language faculty must do this kind of inference at two levels. It must infer both the structures of the sentences produced by speakers and the grammars of the speakers that produce those sentences. Through our discussion, we have seen that the cognitive structure underlying sentences (i.e., the grammar of the language) is highly complex. Our grammatical system includes knowledge of the sounds our language makes use of and the rules governing their distributions, the meanings of words and how they can be combined into sentences, the hierarchical structures of those sentences and dependencies that can hold between elements of that structure, and the ways those structural arrangements give rise to specific sentence interpretations. The architecture of the human language faculty further constrains which rules, structures, and interpretations are possible in any human language. Because children share this cognitive architecture with the rest of the human species, their language faculty is similarly constrained in the types of grammars it can infer. Thus, children's language learning process is shaped not only by their experience with the speech of their community members but by the structure of the language learning mechanism that interprets that experience.

In some ways, our discussion of language acquisition has been idealized: The language learning process can differ for children who are not monolingual, hearing, or typically developing. But these differences are often variations on the same theme. Bilingual children acquire the phonology, lexicon, syntax, and semantics of each language they are exposed to, although they do so with greater ease and proficiency if they hear both languages consistently from an early age. (See Hoff et al., 2012; J. S. Johnson & Newport, 1989; Oyama, 1978; Pearson, Fernandez, & Oller, 1993, among others.) Deaf children exposed to signed languages from an early age acquire a full grammatical system with all of the same components as a spoken language, but one that pairs meaning with visual instead of auditory signals (e.g., Stokoe, 1960). However, in some severe cases, the language learning process can be disrupted by factors intrinsic or extrinsic to the language learner. Cognitive or developmental disabilities can affect children's ability to produce language or process the language they hear, resulting in expressive or receptive language disorders that may persist past childhood (Aram, Ekelman, & Nation, 1984; Bishop, 1997; Clahsen, 1991; Paul, 2007). Children who are deprived of linguistic input until late in development may display grammatical deficiencies into adulthood (Curtiss, 1976; Mayberry & Eichen, 1991; Newport, 1990; Senghas & Coppola, 2001). This finding suggests that the early childhood years are a sensitive period for the development of grammar.

Studying the development of language reveals the complex interaction between children's experience and the tools they bring to this challenging task. Only human children develop language, because only human children are equipped with the cognitive capacities to do so: the capacity to represent complex concepts and understand what other humans mean, to detect patterns in the auditory or visual signals used to convey those meanings, and to interpret those patterns in just the right way to infer the same complex cognitive system as the other language users in their community. They are able to succeed in this task because they are not so different from us. Like other members of the human species, children are equipped with a cognitive faculty specialized for language that guides their process of inferring just the right grammar from their experience. Language acquisition provides us with a window into the rich structure of this human language faculty and how it develops in interaction with its environment and the rest of human cognition.

REFERENCES

Adani, F., Forgiarini, M., Guasti, M. T., & Van der Lely, H. K. (2014). Number dissimilarities facilitate the comprehension of relative clauses in children with (grammatical) specific language impairment. *Journal of Child Language*, *41*(4), 811–841.

Adani, F., Van der Lely, H. K., Forgiarini, M., & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: a comprehension study with Italian children. *Lingua*, *120*(9), 2148–2166.

Allen, S. E. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics*, *38*(3), 483–521.

Aram, D. M., Ekelman, B. L., & Nation, J. E. (1984). Preschoolers with language disorders 10 years later. *Journal of Speech, Language, and Hearing Research*, *27*(2), 232–244.

Arnon, I. (2009). Rethinking child difficulty: the effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language*, *37*(1), 27–57.

Arosio, F., Yatsushiro, K., Forgiarini, M., & Guasti, M. T. (2012). Morphological information and memory resources in children's processing of relative clauses in German. *Language Learning and Development*, *8*(4), 340–364.

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: evidence from 2-year-olds. *Cognition*, *114*(3), 442–446.

Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, *9*(1), 7–31.

Baker, M. C. (1988). *Incorporation: a theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Baker, M. C. (2001). The natures of nonconfigurationality. In R. K. Baltin & C. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (pp. 407–438). Malden, MA: Blackwell.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*(1), 3–26.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 874–890.

Baldwin, D. A. (1993). Early referential understanding: infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*(5), 832.

Baldwin, D. A., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development*, 1915–1939.

Bar-Shalom, E., & Snyder, W. (1997). Root infinitives in child Russian: a comparison with Italian and Polish. In R. Shillcock, N. Sorace, & C. Heycock (Eds.), *Proceedings of the GALA '97 Conference on Language Acquisition: Knowledge Representation and Processing*. Edinburgh, UK: University of Edinburgh.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, *4*(2), 159–219.

Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp. 96–151). Cambridge, MA: Blackwell.

Beckwith, R., Tinker, E., & Bloom, L. (1989). "The acquisition of non-basic sentences." Paper presented at the Boston University Conference on Language Development.

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, *122*(10), 1053–1069.

Bever, T. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.), *Cognition and the Development of Language* (pp. 279–362). New York, NY: Wiley.

Bishop, D. V. M. (1997). *Uncommon Understanding. Development and Language Disorders in Children*. Hove, UK: Psychology Press.

Bloom, L. (1970). *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, *21*(4), 491–504.

Bloom, P. (1994). Possible names: the role of syntax-semantics mappings in the acquisition of nominals. *Lingua*, *92*, 297–329.

Blythe, R. A., Smith, K., & Smith, A. D. (2010). Learning times for large lexicons through cross-situational learning. *Cognitive Science*, *34*(4), 620–642.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (pp. 123–172). Dordrecht, the Netherlands: D. Reidel.

Borer, H., & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language & Linguistic Theory*, *10*(2), 147–189.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychological Science*, *16*(4), 298–304.

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: an investigation of young German- and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics*, *20*(3), 539–570.

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, *61*(1–2), 93–125.

Brooks, P. J., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, *35*(1), 29.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Büring, D. (1997). The great scope inversion conspiracy. *Linguistics and Philosophy*, *20*(2), 175–194.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264–293). Cambridge, MA: MIT Press.

Carey, S. (2009). Where our number concepts come from. *Journal of Philosophy*, *106*(4), 220.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Cartwright, T. A., & Brent, M. R. (1997). Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition*, *63*(2), 121–170.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*(2), 159–199.

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, *30*(3), 348–369.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development*, *10*(1), 1–18.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69–77.

Chien, Y.-C., & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, *1*(3), 225–295.

Chomsky, N. (1957). *Syntactic Structures*. The Hague, the Netherlands: Mouton.

Chomsky, N. (1975). *Reflections on Language*. New York, NY: Pantheon.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht, the Netherlands: Foris.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, *95*(3), 1570–1580.

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, *2*(3), 385–394.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, *51*(1–2), 61–75.

Clahsen, H. (1991). *Child Language and Developmental Dysphasia: Linguistic Studies of the Acquisition of German* (Vol. 2). N.P.: John Benjamins.

Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2009). Equal treatment for all antecedents: how children succeed with Principle B. *Linguistic Inquiry*, *40*(3), 446–486.

Corrêa, L. M. S. (1995). An alternative assessment of children's comprehension of relative clauses. *Journal of Psycholinguistic Research*, *24*(3), 183–203.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers. In D. Dowty, D. Kartunnen, & A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives* (pp. 94–128). Cambridge, UK: Cambridge University Press.

Crain, S., & McKee, C. (1985). The acquisition of structural restrictions on anaphora. In *Proceedings of NELS* (Vol. 15, pp. 94–110). Amherst, MA: Graduate Linguistic Student Association.

Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of "pure reason" in infancy. *Cognition*, *72*(3), 237–267.

Curtiss, S. (1976). *Genie: A Psycholinguistic Study of a Modern-Day Wild Child*. New York, NY: Academic Press.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*(3), 133–142.

de Carvalho, A., Dautriche, I., & Christophe, A. (2016). Preschoolers use phrasal prosody online to constrain syntactic analysis. *Developmental Science*, *19*(2), 235–250.

Dehaene, S. (2009). Origins of mathematical intuitions. *Annals of the New York Academy of Sciences*, *1156*(1), 232–259.

Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, *370*(6487), 292–295.

Demuth, K. (1989). Maturation and the acquisition of the Sesotho passive. *Language*, *65*(1) 56–80.

de Villiers, J., Flusberg, H. B. T., Hakuta, K., & Cohen, M. (1979). Children's comprehension of relative clauses. *Journal of Psycholinguistic Research*, *8*(5), 499–518.

Dollaghan, C. (1985). Child meets WordFast mapping in preschool children. *Journal of Speech, Language, and Hearing Research*, *28*(3), 449–454.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303–306.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition*, *91*(2), 173–190.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences*, *105*(29), 9926–9930.

Feldman, N. H., Myers, E. B., White, K. S., Griffiths, T. L., & Morgan, J. L. (2013). Word-level information influences phonetic learning in adults and infants. *Cognition*, *127*(3), 427–438.

Fillmore, C. J. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in Linguistic Theory*. New York, NY: Holt, Rinehart, & Winston.

Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(2), 143–149.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, *23*(3), 331–392.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1990). Developmental changes in young children's knowledge about the mind. *Cognitive Development*, *5*(1), 1–27.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*(2), 93–126.

Frazier, L., & d'Arcais, G. B. F. (1989). Filler driven parsing: a study of gap filling in Dutch. *Journal of Memory and Language*, *28*(3), 331–344.

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology*, *17*(14), 1208–1211.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*(1), 67–88.

Gagliardi, A., Mease, T. M., & Lidz, J. (2016). Discontinuous development in the acquisition of filler-gap dependencies: evidence from 15- and 20-month-olds. *Language Acquisition*, *23*(3), 1–27.

Gelman, R., & Gallistel, C. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language Development: Language, Thought, and Culture* (Vol. 2, pp. 301–334). Hillsdale, NJ: Erlbaum.

Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, *30*(4), 431–451.

Gerken, L. (1994). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, *21*(3), 565–584.

Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, *51*(3), 237–265.

Gerken, L., Landau, B., & Remez, R. E. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, *26*(2), 204.

Gerken, L., & McIntosh, B. J. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*(3), 448.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules abstract knowledge of word order in early sentence comprehension. *Psychological Science*, *17*(8), 684–691.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23–64.

Gleitman, L. R., Gleitman, H., Landau, B., & Wanner, E. (1988). Where learning begins: initial representations for language learning. *Linguistics: The Cambridge Survey*, *3*, 150–193.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436.

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.

Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*(2), 183–206.

Goodluck, H., & Tavakolian, S. (1982). Competence and processing in children's grammar of relative clauses. *Cognition*, *11*(1), 1–27.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: domain specificity in cognition and culture* (pp. 257–293). New York, NY: Cambridge University Press.

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, *36*(3), 227–254.

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access. II. Infant data. *Journal of Memory and Language*, *51*(4), 548–567.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254–260.

Grimshaw, J. (1981). Form, function and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The Logical Problem of Language Acquisition* (pp. 165–182). Cambridge, MA: MIT Press.

Grodzinsky, Y., & Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry*, *24*(1), 69–101.

Guasti, M. T. (2002). *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press.

Guasti, M. T., Dubugnon, C., Hasan-Shlonsky, S., & Schneitter, M. (1996). Les relatives que nous apprenons [Relatives we learn]. *Rivista Di Grammatica Generativa*, *21*, 107–125.

Guilfoyle, E., & Noonan, M. (1988). “Functional categories and language acquisition.” Paper presented at the 13th Annual Boston University Conference on Language Development.

Gutman, A., Dautriche, I., Crabbé, B., & Christophe, A. (2015). Bootstrapping the syntactic bootstrapper: probabilistic labeling of prosodic phrases. *Language Acquisition*, *22*(3), 285–309.

Hacquard, V. (2014). Bootstrapping attitudes. In T. Snider, S. D’Antonio, & M. Weigand (Eds.), *Semantics and Linguistic Theory* (Vol. 24, pp. 330–352). Proceedings of the 24th Semantics and Linguistic Theory Conference (SALT), May 30–June 1. New York University. Linguistic Society of America.

Haegeman, L. (1995). Root infinitives, tense, and truncated structures in Dutch. *Language Acquisition*, *4*(3), 205–255.

Haendler, Y., Kliegl, R., & Adani, F. (2015). Discourse accessibility constraints in children’s processing of object relative clauses. *Frontiers in Psychology*, *6*.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “Number Sense”: the approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120.

Halberda, J., Taing, L., & Lidz, J. (2008). The development of “most” comprehension and its potential dependence on counting ability in preschoolers. *Language Learning and Development*, *4*(2), 99–121.

Hall, D. G., & Waxman, S. R. (1993). Assumptions about word meaning: individuation and basic-level kinds. *Child Development*, *64*(5), 1550–1570.

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How two- and four-year-old children interpret adjectives and count nouns. *Child Development*, *64*(6), 1651–1664.

Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech*, *51*(1–2), 23–44.

Hamburger, H., & Crain, S. (1982). Relative acquisition. *Language Development*, *1*, 245–274.

Hansen, M. B. (2010). If you know something, say something: young children’s problem with false beliefs. *Frontiers in Psychology*, *1*, 23.

Harrigan, K., Hacquard, V., & Lidz, J. (2016). Syntactic bootstrapping in the acquisition of attitude verbs: think, want and hope. In K.-M. Kim, P. Umbal, T. Block, Q. Chan, T. Cheng, K. Finney, M. Katz . . . L. Shorten, *Proceedings of the 33rd West Coast Conference on Formal Linguistics* (pp. 196–206). Somerville, MA: Cascadilla Proceedings Project.

Harris, F. N., & Flora, J. A. (1982). Children’s use of get passives. *Journal of Psycholinguistic Research*, *11*(4), 297–311.

Harris, T., & Wexler, K. (1996). The optional-infinitive stage in child English: evidence from negation. In H. Clahsen (Ed.), *Language Acquisition and Language Disorders* (Vol. 14, p. 1). Amsterdam, the Netherlands: John Benjamins.

He, A. X., & Lidz, J. (2017). Development of the verb-event link in 14- and 18-month-old English-learning infants. *Journal of Language Learning and Development*, *13*(3), 335–356.

He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, *30*(1), 14–29.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: an examination of fast mapping. *Child Development*, 1021–1034.

Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Sciences*, *18*(4), 167–170.

Hicks, J., Maye, J., & Lidz, J. (2007). "The role of function words in infants' syntactic categorization of novel words." Paper presented at the Linguistic Society of America Annual Meeting, Anaheim, CA.

Higginbotham, J., & May, R. (1981). Questions, quantifiers and crossing. *Linguistic Review*, *1*(1), 41–80.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for Assessing Children's Syntax* (pp. 105–124). Cambridge, MA: MIT Press.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27.

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*(2), 122–127.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: the role of determiners in the syntactic categorization of lexical elements. *Infancy*, *5*(3), 341–353.

Horn, L. (1989). *A Natural History of Negation*. Chicago, IL: University of Chicago Press.

Huang, Y. T., Zheng, X., Meng, X., & Snedeker, J. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language*, *69*(4).

Hunter, T., & Lidz, J. (2013). Conservativity and learnability of determiners. *Journal of Semantics*, *30*(3), 315–334.

Hyams, N. (1986). *Language Acquisition and the Theory of Parameters*. Dordrecht, the Netherlands: D. Reidel.

Hyams, N. (1992). A reanalysis of null subjects in child language. In J. Weissborn, H. Goodluck, & T. Roeper (Eds.), *Theoretical Issues in Language Acquisition: Continuity and Change in Development* (pp. 249–267). Hillsdale, NJ: Psychology Press.

Hyams, N. (2011). Missing subjects in early child language. In J. De Villiers & T. Roeper (Eds.), *Handbook of Generative Approaches to Language Acquisition* (pp. 13–52). Dordrecht: Springer.

Hyams, N., & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, *24*(3), 421–459.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, *13*(2), 339–345.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*(1), 1–23.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*(2), 252–293.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*(5), 630–645.

Kazanina, N., & Phillips, C. (2001). Coreference in child Russian: distinguishing syntactic and discourse constraints. In *Proceedings of the 25th*

Boston University Conference on Language Development. Boston, MA: Cascadilla Press.

Kedar, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development*, *77*(2), 325–338.

Keenan, E. L., & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, *9*(3), 253–326.

Kidd, E., & Bavin, E. L. (2002). English-speaking children's comprehension of relative clauses: evidence for general-cognitive and language-specific constraints on development. *Journal of Psycholinguistic Research*, *31*(6), 599–617.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: a cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860–897.

Kim, Y.-J. (2000). Subject/object drop in the acquisition of Korean: a cross-linguistic comparison. *Journal of East Asian Linguistics*, *9*(4), 325–351.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190*(4209), 69–72.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), F13–F21.

Labelle, M. (1990). Predication, wh-movement, and the development of relative clauses. *Language Acquisition*, *1*(1), 95–119.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: evidence from the blind child*. Cambridge, MA: Harvard University Press.

Lasnik, H. (1976). Remarks on coreference. *Linguistic Analysis*, *2*, 1–22.

Lasnik, H. (1989). On certain substitutes for negative data. In R. J. Matthews & W. Demopoulos (Eds.), *Learnability and Linguistic Theory* (pp. 89–105). Dordrecht, the Netherlands: Kluwer.

Legate, J. A., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, *14*(3), 315–344.

Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal Cognition: A Multidisciplinary Debate* (pp. 121–141). Oxford, UK: Oxford University Press.

Lewis, S., Hacquard, V., & Lidz, J. (2017, March). "Think" pragmatically: children's interpretation of belief reports. *Language Learning and Development*, 1–23.

Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: a reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language*, *32*(2), 169.

Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, *29*(1), 117–123.

Lidz, J. (2006). The grammar of accusative case in Kannada. *Language*, *82*(1), 10–32.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition*, *87*(3), 151–178.

Lidz, J., & Musolino, J. (2002). Children's command of quantification. *Cognition*, *84*(2), 113–154.

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months. *Cognition*, *89*(3), 295–303.

Locke, J. (1998). *An Essay Concerning Human Understanding*. (R. Woolhouse, Ed.) (Reprint edition). London, UK: Penguin Classics. (Originally published 1690.)

Lukyanenko, C., Conroy, A., & Lidz, J. (2014). Is she patting Katie? Constraints on pronominal reference in 30-month-olds. *Language Learning and Development*, *10*(4), 328–344.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology*, *58*(4), 441–486.

Maratsos, M. P., & Abramovitch, R. (1975). How children understand full, truncated, and anomalous passives. *Journal of Verbal Learning and Verbal Behavior*, *14*(2), 145–157.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, *92*, 199–227.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cognitive Psychology*, *16*(1), 1–27.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, *78*(2), 91–121.

Mayberry, R. I., & Eichen, E. B. (1991). The long-lasting advantage of learning sign language in childhood: another look at the critical period for language acquisition. *Journal of Memory and Language*, *30*(4), 486–512.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, *11*(1), 122–134.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111.

McKee, C., McDaniel, D., & Snedeker, J. (1998). Relatives children say. *Journal of Psycholinguistic Research*, *27*(5), 573–596.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), 33–42.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*(22), 9014–9019.

Mersad, K., & Nazzi, T. (2012). When Mommy comes to the rescue of statistics: infants combine top-down and bottom-up cues to segment speech. *Language Learning and Development*, *8*(3), 303–315.

Mesgarani, N., David, S. V., Fritz, J. B., & Shamma, S. A. (2008). Phoneme representation and classification in primary auditory cortex. *Journal of the Acoustical Society of America. 123*(2), 899–909.

Millotte, S., Morgan, J., Margules, S., Bernal, S., Dutat, M., & Christophe, A. (2010). Phrasal prosody constrains word segmentation in French 16-month-olds. *Journal of Portuguese Linguistics*, *9*(2), 10.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91–117.

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: the incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267–293.

Mintz, T. H., Newport, E. L., & Bever, T. G. (1995). Distributional regularities of form class in speech to young children. In J. N. Beckman (Ed.), *Proceedings of NELS 25* (Vol. 2, pp. 43–54). Amherst, MA: Graduate Linguistic Student Association.

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26*(4), 393–424.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, *96*(2), 143–182.

Morgan, J. L. (1986). *From Simple Input to Complex Grammar*. Cambridge, MA: MIT Press.

Morgan, J. L., & Demuth, K. (1996). Signal to syntax: an overview. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, pp. 1–22. Mahwah, NJ: Erlbaum.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, *19*(4), 498–550.

Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, *20*(1), 67–85.

Musolino, J., Crain, S., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, *38*(1), 1–32.

Musolino, J., & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics*, *44*(4).

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357–374.

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407–420.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756–766.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology*, *42*(4), 643–655.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28.

Noble, C. H., Rowland, C. F., & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive Science*, *35*(5), 963–982.

Odic, D., Halberda, J., Pietroski, P., & Lidz, J. (n.d.). "The language-number interface: evidence from the acquisition of most." Manuscript submitted for publication.

Odic, D., Pietroski, P., Hunter, T., Lidz, J., & Halberda, J. (2013). Young children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 451–461.

Omaki, A., Davidson White, I., Goro, T., Lidz, J., & Phillips, C. (2014). No fear of commitment: children's incremental interpretation in English and Japanese wh-questions. *Language Learning and Development*, *10*(3), 206–233.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), 13–23.

Oyama, S. (1978). The sensitive period and comprehension of speech. *NABE Journal*, *3*(1), 25–40.

Paul, R. (2007). *Language Disorders from Infancy through Adolescence: Assessment & Intervention*. N.P.: Elsevier Health Sciences.

Pearson, B. Z., Fernandez, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: comparison to monolingual norms. *Language Learning*, *43*(1), 93–120.

Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Phillips, C. (1995). Syntax at age two: cross-linguistic differences. *MIT Working Papers in Linguistics*, *26*, 325–382.

Pine, J. M., & Lieven, E. V. (1997). Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics*, *18*(2), 123–138.

Pine, J. M., & Martindale, H. (1996). Syntactic categories in the speech of young children: the case of the determiner. *Journal of Child Language*, *23*(2), 369–395.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Platzack, C. (1990). A grammar without functional categories: a syntactic study of early Swedish child language. *Nordic Journal of Linguistics*, *13*(2), 107–126.

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language*, *69*(1), 1–33.

Polka, L., & Werker, J. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421.

Quine, W. V. O. (2013). *Word and Object*. Cambridge, MA: MIT Press. (Originally published 1960.)

Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Cambridge, MA: Blackwell.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349–351.

Redington, M., & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: a distributional perspective. *Language and Cognitive Processes*, *13*(2–3), 129–191.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: a powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*(4), 425–469.

Reinhart, T. (1981). Definite NP anaphora and c-command domains. *Linguistic Inquiry*, *12*(4), 605–635.

Rizzi, L. (1993). Some notes on linguistic theory and language development: the case of root infinitives. *Language Acquisition*, *3*(4), 371–393.

Rizzi, L. (1994). Early null subjects and root null subjects. In T. Hoekstra & B. D. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 151–176). Amsterdam, the Netherlands: John Benjamins.

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*(1), 27–33.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition*, *69*(2), 105–134.

Schaeffer, J., & Ben Shalom, D. (2004). On root infinitives in child Hebrew. *Language Acquisition*, *12*(1), 83–96.

Seidl, A., Cristià, A., Bernard, A., & Onishi, K. H. (2009). Allophonic and phonemic contrasts in infants' learning of sound patterns. *Language Learning and Development*, *5*(3), 191–202.

Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object wh-questions. *Infancy*, *4*(3), 423–436.

Senghas, A., & Coppola, M. (2001). Children creating language: how Nicaraguan sign language acquired a spatial grammar. *Psychological Science*, *12*(4), 323–328.

Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics*, *26*(3), 437–462.

Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research*, *41*(4), 874–886.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior*, *13*(3), 272–281.

Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, *119*(6), EL61–EL67.

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, *11*(3), 407–413.

Shi, R., Marquis, A., Gauthier, B., Bamman, D., Magnitskaia, T., & Zaller, C. (2006). Segmentation and representation of function words in preverbal French-learning infants. In *Proceedings of the 30th Annual Boston University Conference on Language Development* (Vol. 2, pp. 549–560). Somerville, MA: Cascadilla Press.

Shi, R., & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, *15*(5), 517–533.

Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, *25*(1), 169–201.

Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*(2), 187–198.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21.

Shipley, E. F., Smith, C. S., & Gleitman, L. R. (1969). A study in the acquisition of language: free responses to commands. *Language*, 322–342.

Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Developmental Psychology*, *28*(2), 273.

Smith, L. [B.], & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.

Stokoe, W. C. (1960). *Sign language structure: an outline of the visual communication systems of the American deaf*. Buffalo, NY: Department of Anthropology and Linguistics, University of Buffalo.

Stromswold, K., Eisenband, J., Norland, E., & Ratzan, J. (2002). Tracking the acquisition and processing of English passives: using acoustic cues to disambiguate actives and passives. In *CUNY Conference on Sentence Processing* (Vol. 21–23). New York, NY.

Sutton, M. (2015). *Competence and performance in the development of principle C*. Doctoral dissertation, University of Maryland, College Park, MD.

Sutton, M., Fetters, M., & Lidz, J. (2012). Parsing for principle C at 30 months. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds.), *Proceedings of the 36th Boston University Conference on Language Development* (pp. 581–593). Somerville, MA: Cascadilla Press.

Takahashi, E. (2009). Beyond statistical learning in the acquisition of phrase structure. Doctoral dissertation, University of Maryland, College Park, MD.

Takahashi, E., & Lidz, J. (2008). Beyond statistical learning in syntax. In A. Gavarró & M. João Freitas, *Proceedings of Generative Approaches to Language Acquisition* (pp. 444–445). Cambridge, UK: Cambridge Scholars Publishing.

Tavakolian, S. (1981). The conjoined-clause analysis of relative clauses. *Language Acquisition and Linguistic Theory*, (pp. 167–187). Cambridge, MA: MIT Press.

Taylor, M., & Gelman, S. A. (1988). Adjectives and nouns: children's strategies for learning new words. *Child Development*, *59*, 411–419.

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: the role of transitional probability. *Language Learning and Development*, *3*(1), 1–42.

Thornton, R., & Wexler, K. (1999). *Principle B, VP Ellipsis, and Interpretation in Child Grammar*. Cambridge, MA: MIT Press.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*(3), 209–253.

Tomasello, M. (2003). *Constructing a Language: a Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Turner, E. A., & Rommetveit, R. (1967). The acquisition of sentence voice and reversibility. *Child Development*, *38*, 649–660.

Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, *22*(4), 562.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1–2), 21–81.

Valian, V., & Aubry, S. (2005). When opportunity knocks twice: two-year-olds' repetition of sentence subjects. *Journal of Child Language*, *32*(3), 617–641.

Valian, V., & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language*, *23*(1), 103–128.

Valian, V., Hoeffner, J., & Aubry, S. (1996). Young children's imitation of sentence subjects: evidence of processing limitations. *Developmental Psychology*, *32*(1), 153.

Valian, V., Solt, S., & Stewart, J. (2009). Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language*, *36*(4), 743–778.

Viau, J., Lidz, J., & Musolino, J. (2010). Priming of abstract logical representations in 4-year-olds. *Language Acquisition*, *17*(1–2), 26–50.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*(2), 729–742.

Wagers, M. W., & Phillips, C. (2014). Going the distance: memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology*, *67*(7), 1274–1304.

Wang, Q., Lillo-Martin, D., Best, C. T., & Levitt, A. (1992). Null subject versus null object: some evidence from the acquisition of Chinese and English. *Language Acquisition*, *2*(3), 221–254.

Wanner, E., & Gleitman, L. R. (1982). *Language Acquisition: The State of the Art*. Cambridge, UK: Cambridge University Press.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition*, *70*(3), B35–B50.

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, *43*(3), 217–242.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*(3), 257–302.

Waxman, S. R., & Markow, D. B. (1998). Object properties and object kind: twenty-one-month-old infants' extension of novel adjectives. *Child Development*, *69*(5), 1313–1329.

Wellwood, A., Gagliardi, A., & Lidz, J. (2016). Syntactic and lexical inference in the acquisition of novel superlatives. *Language Learning and Development*, *12*(3), 262–279.

Werker, J. F. (1995). Exploring developmental changes in cross-language speech perception. In L. Gleitman, M. Liberman, & D. N. Osherson (Eds.), *An Invitation to Cognitive Science: Language* (2nd ed., Vol. 1, pp. 87–106). Cambridge, MA: MIT Press.

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*(1), 349–355.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental Psychology*, *24*(5), 672–683.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63.

Weverink, M. (1989). "The subject in relation to inflection in child language." Master's thesis, University of Utrecht.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot & N. Hornstein (Eds.), *Verb Movement* (pp. 305–350). Cambridge, UK: Cambridge University Press.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua*, *106*(1), 23–79.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: the psychophysics of number representation. *Psychological Science*, *10*(2), 130–137.

White, A. S., Hacquard, V., & Lidz, J. (2017). Main clause syntax and the labeling problem in syntactic bootstrapping.

Williams, A. (2015). *Arguments in Syntax and Semantics*. Cambridge, UK: Cambridge University Press.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34.

Woodward, A. L., & Markman, E. M. (1998). Early word learning. In W. Damon (Ed.), *Handbook of Child Psychology: Cognition, Perception, and Language* (Vol. 2, pp. 371–420). New York, NY: Wiley.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272.

Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6324–6327.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: simple structural cues to verb meaning. *Child Development*, *83*(4), 1382–1399.

CHAPTER 4

Development of Episodic Memory: Processes and Implications

CHRISTINE COUGHLIN, SARAH LECKEY, AND SIMONA GHETTI

One of my favorite memories is when my best friend and I had hiked to the very top of a mountain in Colorado and we sat there marveling at the beauty of it all. She told me to look around and then close my eyes to take a mental picture. I did. I remember being so in tune to myself and the world around me for one of the first times in my young life. I was 19. I'll never forget this memory with my lovely friend.

—Internet source sharing a favorite memory

Episodic memory is the ability to remember specific moments in time in rich contextual detail. The quote at the beginning of this chapter from an anonymous internet user illustrates the universality of this human experience: Episodic memory involves a mental picture of the event that includes contextual details, such as what the individual was seeing, feeling, and thinking as the event unfolded (Tulving, 1984, 2005). For this reason, it is sometimes referred to as mental time travel, allowing us to travel back in time to mentally relive something that has already happened to us (Suddendorf & Corballis, 2007).

The ability to remember past experiences in detail serves several important functions. For example, research has shown that episodic memory contributes to learning, benefiting reading comprehension (Mirandola, Del Prete, Ghetti, & Cornoldi, 2011), novel inference making (Zeithamova, Schlichting, & Preston, 2012), and the acquisition of semantic knowledge (Gardiner, Brandt, Baddeley, Vargha-Khadem, & Mishkin, 2008; Martins, Guillery-Girard, Jambaqué, Dulac, & Eustache, 2006). Episodic memory also contributes to autobiographical memory and the formation of self-concepts (Nelson & Fivush, 2004). These contributions may extend to the "future self": There is evidence that we rely on content from episodic memories when constructing possible personal future events (Schacter, Addis, & Buckner, 2007). Thus, episodic memory potentially allows individuals to learn from their past and prepare for their future.

Given the important role of episodic memory in our daily lives, and considering how much time the average human spends reminiscing, it is not surprising that a large body of research has been dedicated to elucidating the cognitive and neural processes supporting the development of this ability. In this chapter, we provide an overview of this research, summarizing what is known about the development of this critical ability from infancy to adolescence.

ROLE OF BINDING AND CONTROL PROCESSES

Episodic memory is a rich construct that emerges from the contribution of a number of processes. Despite differences among models of episodic memory (Nadel & Moscovitch, 1997; Shing, Werkle-Bergner, Li, & Lindenberger, 2008; Squire, 1992; Tulving, 1985), there is some convergence on the fact that at least two classes of processes contribute to adult performance (Koriat & Goldsmith, 1996; Ranganath, 2010) and development (Ghetti & Bunge, 2012), namely binding and control processes. Binding processes (also referred to as associative processes) integrate the discrete features of an event (e.g., spatial and temporal information, sensory input, cognitions, and emotions) to create a "bound" mental representation of the event; these processes are thought to be critically supported by the hippocampus (HPC) and surrounding medial temporal cortices (e.g., Eichenbaum, Yonelinas, & Ranganath, 2007; Konkel & Cohen, 2009; Wixted & Squire, 2011). Control processes may promote binding processes by allocating attention and cognitive resources to the encoding of events (Ghetti Lyons & DeMaster, 2012). Or they may act on the bound representation of an event during retrieval via monitoring assessments of accuracy (e.g., How clear is this memory? How confident am I that it is accurate?) or effort (e.g., How difficult was it to retrieve it?). Metacognitive acts change remembering from being a reflexive act, to being an intentional, goal-directed, and reflective activity (Moscovitch, 1992). These processes are thought to be largely supported by cortical mechanisms in the prefrontal and parietal lobes (Ranganath, 2010).

Traditionally, developmental improvements in episodic memory during young childhood were attributed to improvements in binding processes supported by maturation of the HPC, whereas later improvements were attributed to control processes supported largely by the prefrontal cortex (PFC) (e.g., Shing et al., 2008). This view, based on indirect evidence, makes intuitive sense. Research examining memory strategies and metacognition consistently showed difficulties in early childhood and protracted improvement into late childhood (Ghetti, Castelli, & Lyons, 2010; Roebers, 2002; Schneider & Lockl, 2002), implying that early memory functioning likely does not benefit from these processes to a great extent. In contrast, evidence from nonhuman animals indicated maturation of the HPC at times corresponding to infancy or early childhood (e.g., Seress & Ribak, 1995), suggesting that improvements in binding processes primarily contribute to the emergence and early development of memory.

However, research over the past decade has challenged this view. This research has begun to show evidence of continued development of binding processes well beyond early childhood and into adolescence (DeMaster & Ghetti, 2013; Ghetti et al. 2010; Sastre, Wendelken, Lee, Bunge, & Ghetti, 2016; Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2016), as well as evidence of control processes in early childhood (Ghetti, Hembacher, & Coughlin, 2013). Together, this research suggests that episodic memory development likely results from an interaction and refinement of both binding and control processes from infancy, across childhood, and into adolescence. Thus, in the following sections, we consider both types of processes while providing of an overview of episodic memory development during these periods. We note that although episodic memory is supported by numerous cortical regions and subcortical structures (Cabeza & St. Jacques, 2007; Hassabis, Kumaran, & Maguire, 2007; Rugg, Otten, & Henson, 2002; Wixted & Squire, 2011), our discussion of neural underpinnings focuses on the HPC and PFC given that these areas have received the most attention in developmental cognitive neuroscience approaches to episodic memory.

DEVELOPMENT OF EPISODIC MEMORY

Infancy

Infants demonstrate memory capabilities from a very young age (e.g., for their mother's face, voice, and breast milk; Bushnell, Sai, & Mullin, 1989; DeCasper & Fifer, 1980; MacFarlane, 1975). The scientific demonstration of these skills has relied on a number of experimental paradigms assessing memory independent of receptive or expressive language ability, either of which is typically relied on in tasks with verbal participants. In these paradigms, infants are exposed to a to-be-remembered stimulus and then, following a delay, are placed in situations in which they can behave in a manner that indicates memory of the stimulus. Behavior indicative of memory retention might include a visual preference (visual paired comparison tasks; Fantz, 1958), performing previously reinforced actions at an increased rate (conjugate reinforcement; Rovee-Collier & Gekoski, 1979), or imitating an action series that had been previously demonstrated (deferred imitation; see Hayne, 2007, for a review).

Binding Processes

Although there is some disagreement on the extent to which experimental paradigms used in infancy can assess declarative forms of memory requiring conscious thought (e.g., Bauer, Deboer, & Lukowski, 2007; Rovee-Collier, 1997; Schacter & Moscovitch, 1984), it is generally agreed that research using these paradigms has provided important insight into the early development of the capacity to form memory representations that combine elements of unique experiences. From this perspective, irrespective of how memory is expressed (ranging from intentional actions that could be available, to introspection, to automatic actions without awareness), this research can illuminate early binding capacities. For example, Richmond and colleagues (Richmond & Nelson, 2009; Richmond & Power, 2014) provided evidence of early binding processes across two studies using a visual paired comparison task. In these studies, 6-, 9-, and 12-month-olds were familiarized with face-scene pairs and then later shown each scene overlaid with three familiar faces (one of which was originally paired with the scene). Results showed that 6- and 9-month-olds (but not 12-month-olds) preferentially looked toward the face that was originally paired with the scene, suggesting that even 6-month-olds can encode and retrieve item-item associations in some circumstances, and that some reorganization of memory representations may occur toward the end of the first year of life. Experiments utilizing deferred imitation paradigms also suggest that rudimentary binding processes, demonstrated by retention of object-action relations or retention of the temporal order of the actions, may be in place fairly early on. For example, after being shown target actions on specific objects (six times each), both 6- and 9-month-olds successfully reproduced the target actions on the specific objects after a 24-hour delay (Barr, Dowden, & Hayne, 1996). Other work has highlighted developmental improvements across infancy. In a study with 13-, 16-, and 20-month-olds, Bauer and Leventon (2013) found that 16- and 20-month-olds could demonstrate an action sequence that had been demonstrated only once after a 1-month delay, but that 13-month-olds required multiple original exposures to do so. And, 16- and

20-month-olds both required multiple original exposures to demonstrate retrieval after a 3-month delay, whereas 13-month-olds could not do so regardless of the number of original exposures. Interestingly, when subsamples of 13-, 16-, and 20-month-old infants matched for their immediate imitation scores (taken as evidence of comparable encoding ability) were tested after a several-month delay, younger infants demonstrated greater forgetting (Bauer, 2005). This latter finding aligns with other work suggesting that post-encoding consolidation processes may contribute to age differences in memory retention in infancy (Bauer & Larkina, 2014; Pathman & Bauer, 2013).

Overall, these varying effects of exposure and delay on performance indicate that developmental changes in encoding, storage, and retrieval occur across the first 2 years of life. Indeed, despite evidence of striking memory skills in even young infants, some of the most important signatures of binding capacity are not reliably demonstrated until the end of infancy—namely, the demonstration of memory for associations between items and their location in space when no cues are present (Newcombe, Balcomb, Ferrara, Hansen, & Koski, 2014), as well as the temporal order of actions that do not have enabling or known relations (Bauer, Hertsgaard, Dropik, & Daly, 1998).

Altogether, these and other behavioral findings (for reviews, see Bauer, San Souci, & Pathman, 2010; Hayne, 2007; Mullally & Maguire, 2014; Richmond & Nelson, 2007) suggest that improvements in binding processes contribute to the development of episodic memory across the first 2 years of life. This notion aligns with what is known about early maturational change in the HPC, the neural structure that supports binding processes (Konkel & Cohen, 2009). Although substantial structural changes occur prenatally (Seress, Ábrahám, Tornóczky, & Kosztolányi, 2001), hippocampal volume nearly doubles during the first year of life and continues to increase rapidly during the second year as well (Gilmore et al., 2012). Histological studies suggest that these volumetric changes are especially apparent in the dentate gyrus and Cornu Ammonis subfield 3 (CA3; Insausti et al., 2010; Lavenex & Banta Lavenex, 2013; Seress et al., 2001), hippocampal subfields that may be particularly important for developmental improvements in the encoding and retrieval of complex associations. (See Lee, Johnson, & Ghetti, 2017, for a review.) Interestingly, one consequence of this rapid hippocampal development may be infantile amnesia (i.e., adults' and older children's inability to remember events that happened to them early in life; Bauer 2007; Howe & Courage, 1993; Peterson, 2002; Rubin, 2000). Indeed, Josselyn and Frankland (2012) have suggested that hippocampal neurogenesis during this period may replace existing synaptic connections and thus interfere with the long-term memory storage of early life events. Based on this proposal, although hippocampal neurogenesis confers new abilities to encode and retain complex event information for newly encountered events, it may also contribute to memory loss for events that preceded the neurogenesis. Although compelling evidence of a relation between neurogenesis and memory loss has been reported with rodent models (Akers et al., 2014), evidence from humans is currently lacking. The use of neuroimaging techniques with increasingly younger populations may enable researchers to connect volumetric and functional changes in the HPC with memory retention and loss during infancy and early childhood.

Control Processes

Thus far our discussion of the building blocks of episodic memory during infancy has focused on binding processes without any mention of control processes. Assessing the presence or absence of control processes in infants presents significant difficulties. This is because the overt behaviors used to assess infant memory do not provide insight into infants' cognitions about their memory state (e.g., whether an infant is aware of task difficulty or confident in the retrieved memory). Despite this limitation, nascent work suggests that precursors to memory control processes may emerge during the first 2 years of life. For example, Goupil, Romand-Monnier, and Kouider (2016) have shown that 20-month-olds selectively seek help when faced with difficult memory decisions in order to improve their performance. Although this study cannot rule out the possibility that infants' bids for help, signaled by turning their heads toward their parents, reflect a learned association between undesirable outcomes (e.g., goal thwarting as in failure to find an object) and their parents' intervention, it is possible that infants may possess a rudimentary capacity to monitor memory-based difficulties and act strategically to overcome such difficulties. Indeed, this is consistent with early work showing that infants of similar age look more frequently toward a location where an object is hidden if they are told that they are later responsible for retrieving it compared to if they are told that someone else is responsible for retrieving it, or if the object is in plain view, suggesting the presence of precursors of rehearsal-like strategies (DeLoache, Cassidy, & Brown, 1985). These early behaviors may be limited in scope, but they might lay the foundation for later development. Other work has shown that executive control relates positively to 15- and 20-month-olds' performance on a sequence imitation task with built-in interference (Wiebe, Lukowski, & Bauer, 2010), and that infants use communicative gestures to facilitate their learning (Begus, Gliga, & Southgate, 2014). These data converge on the idea that self-regulation, an important component of control processes, might tangibly influence even infants' memory-related behavior. These findings also align with work showing that the PFC is functionally involved in numerous cognitions during infancy (Grossmann, 2013), and that cortical change in this region occurs during infancy (Li et al., 2014).

Overall, the literature on infant development has produced a wealth of evidence on the early capacity to bind information in representations that capture unique combinations of event-space-time information, which is necessary for episodic memory. In addition, the literature begins to provide some insight on the early development of processes that might guide the encoding and retrieval of detailed memories in infancy.

Early Childhood

The rapid development of episodic memory during infancy is followed by continued developmental change during early childhood. Most children now possess sufficient language skills (Owens, 1984) to allow experimenters to assess their memory states verbally, in addition to using paradigms in which memory is reflected in behaviors. Studies using these methods have led researchers to attribute continued developmental changes to improvements in both binding and control processes.

Binding Processes

Improvements in binding processes during early childhood are evident from children's increasing ability to report contextual information about particular events, such as where the event occurred (item-space association), when the event occurred (item-time association), and with what other events it co-occurred (item-item association) (Lee, Wendelken, Bunge, & Ghetti, 2016).

Sluzenski and colleagues (Sluzenski, Newcombe, & Kovacs, 2006) have demonstrated gradual developmental change in 4- to 6-year-olds' ability to remember these associations. In their study, children were shown sets of pictures that included a picture of an animal, a picture of a background, and a picture of the animal superimposed onto the background. During a subsequent retrieval phase, memory for the individual animals, individual backgrounds, and animal/background combinations was tested. Results indicated significant age-related improvements in memory for the animal/background combinations, the condition that placed highest demands on binding processes, but not memory for the individual animals or backgrounds. Work by Lloyd, Doydum, and Newcombe (2009) suggested that these improvements may be particularly due to improvements in children's ability to retrieve bound associations. In their study, 4- and 6-year-olds were tested on their ability to remember objects, backgrounds, and object/background associations during either an immediate working memory test or a delayed memory test. Age-related improvements in the object/background condition were observed during the delayed memory test but not during the working memory test, pointing toward retrieval-related deficits for younger children. Other work by Bauer and colleagues (Bauer, Doydum, Pathman, Larkina, Güler, & Burch, 2012) lent further support for the development of binding processes during this period. In a study with 4-, 6-, and 8-year-olds, they observed age-related improvement in memory for specific laboratory events, but even greater improvement in memory for events in conjunction with their unique locations (i.e., memory of the item-space association).

Neuroimaging work adds a layer of complexity to these described behavioral findings. Given that the HPC supports binding processes, one might predict an increase in overall hippocampal volume during this period. Results from several studies suggest that this may not be the case. Although Uematsu et al. (2012) documented increases in overall hippocampal volume from infancy into childhood, Lee et al. (2015) found no increases between the ages of 2 and 4 years, and Gogtay et al. (2006) found no increases between the ages of 4 to 25 years. Results from these studies may differ for a variety of reasons (e.g., differences in developmental window, sample size, and neuroimaging methods), but they do suggest that potential increases in overall hippocampal volume are likely to be modest at best and that other indices of neural change should be considered. Recent work by Riggins and colleagues (Riggins, Blankenship, & Mulligan, 2015) supported this view. Using structural scans from 4- and 6-year-olds, they found evidence of marginal increases in the volume of the left hippocampal body and hippocampal tail bilaterally. Of greater importance, they found a positive association between the volume of the hippocampal head and source memory performance in 6-year-olds but not in 4-year-olds. Given that no age differences in hippocampal head volume were observed, this finding indicates developmental change in hippocampal function in the absence of volumetric change. Additional work by Riggins, Geng, Blankenship, and Redcay (2016) examined differences in anterior and posterior hippocampal resting-state functional connectivity in 4- and 6-year-olds. Results revealed age-related differences in functional connectivity, as well as in the relation between functional connectivity and memory performance; increased connectivity within an anterior/posterior hippocampal network tended to relate to increased memory performance in 6-year-olds, whereas functional connectivity between the HPC and other regions (e.g., superior temporal gyrus and middle temporal gyrus) tended to relate to increased memory performance in 4-year-olds. Altogether, these results suggest that hippocampal change in early childhood contributes to the development of binding processes and highlight the importance of considering the differential roles of hippocampal subregions and subfields.

Control Processes

Just as early childhood is a time for improvements in binding processes, so too is it a time for improvements in control processes. Although research traditionally emphasized children's limitations in the capacity to exert control over memory processes due to a limited understanding of memory functioning (Kreutzer, Leonard, & Flavell, 1975; Wellman, 1977) or the origin of mental representations (e.g., Perner, Kloo & Stottinger, 2007), these limitations may also have reflected the use of experimental paradigms that were inappropriate for testing young children. Over the past decade, an increased effort to adopt child-friendly procedures (e.g., paradigms that avoid heavy verbal demands, are simple, and are based in familiar contexts; Wellman, 1988) has yielded impressive evidence of control processes in children as young as three. For example, Balcomb and Gerken (2008) had 3.5-year-olds complete two recognition memory tests examining their memory for paired associates (animals paired with objects). Children were allowed to skip trials in the first test but were required to answer all trials during the second test. Results showed that children were more likely to skip trials in the first test that they subsequently got wrong in the second test compared to items that they subsequently got correct. Thus, when given the opportunity to do so, children chose to skip trials for which their memories were least accurate. Later work by Hembacher and Ghetti (2014) demonstrated that 3- to 5-year-olds not only withhold memory decisions that are less accurate, but also report being less confident in their withheld memory decisions. In their study, children encoded pictures of objects and were then asked to report which objects they had seen before during a retrieval phase. Critically, the children were also asked to report how confident they were in each memory decision and to choose whether or not they wanted each memory decision to be evaluated for a possible reward. Although there were no age-related differences in the accuracy of memory judgments, older children were more confident for correct versus incorrect responses, and were also more apt to exclude their least-confident memories (Figure 4.1) which resulted in greater accuracy for selected memories.

Thus, despite the comparable accuracy of their actual memories, older children's memory performance was better because they were more apt to exclude their least-confident and least-accurate responses to be evaluated for a potential reward. Results therefore demonstrate age-related improvements in both memory monitoring and the strategic regulation of memory responses during early childhood.

Other work provides additional support for improvements in strategy use during this time, while also indicating that strategy use does not necessarily translate to gains in recall performance. In one study, 4- to 6-year-olds were given a set of items to play with and then later tested on their memory for a subset of the items (Baker-Ward, Ornstein, & Holden, 1984). Some of the children were assigned to a memory condition in which they were instructed to memorize a subset of the items; others were assigned to control conditions that did not include any memorization instructions. Children's interaction with the toys differed as a function of whether they were assigned to the memory or control conditions, and age-related increases

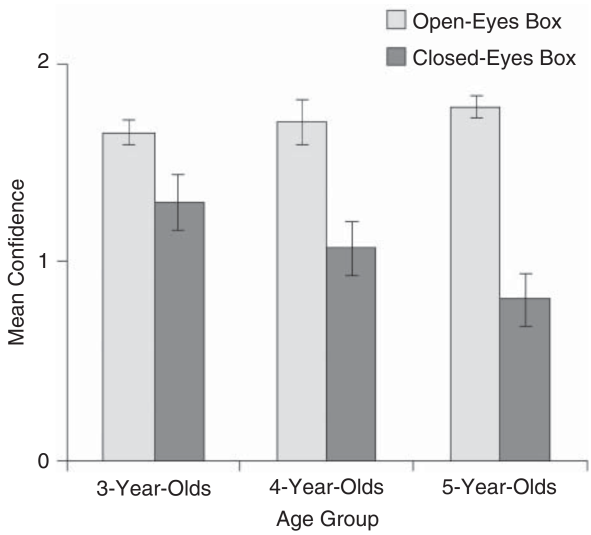


Figure 4.1 This figure presents mean confidence ratings for 3- to 5-year-olds' memory decisions as a function of whether they chose to include (i.e., put in open-eyes box) or exclude (i.e., put in closed-eyes box) their memory decision to be evaluated for a potential prize based on overall performance. Only trials with accurate responses are shown, and error bars represent standard errors. Older children were more apt to exclude their least confident memories.
SOURCE: Adapted from Hembacher & Ghetti (2014).

in the visual inspection and naming of the objects were observed in the memory group. Of importance, though, only 6-year-olds demonstrated greater recall when assigned to the memory versus control conditions. In a related study, Newman (1990) found that 4- to 5-year-olds recalled more toys following a play condition in which they were instructed "to play" versus a recall condition in which they were instructed "to remember." Miller and Seier (1994) have labeled younger children's lack of recall improvement despite having engaged in spontaneous strategy use a "utilization deficiency." These researchers stressed that rudimentary strategy production does not always translate to performance gain or preclude further development of that strategy. They also pointed out possible developmental differences in the level of spontaneity and effort associated with strategy use with age. These differences may be due to developmental improvements in working memory (Gathercole, 1998; 1999; Luciana & Nelson, 1998) and inhibitory control (Diamond, 2002) during early childhood. Thus, although young children demonstrate and sometimes benefit from memory strategies during early childhood, significant development continues to occur with age.

Although the described age-related improvements in control processes are likely due to cortical changes within the PFC, very few studies have examined this region during early childhood. (See Tsujimoto, 2008, for review.) This lack of research is primarily due to difficulties using neuroimaging methods with young children. Nonetheless, there is some evidence of increases in the gray matter of the PFC during early childhood, in addition to later developmental change (Giedd et al., 1999). And postmortem studies indicate substantial age-related changes in the neuronal density, synaptic density, and cellular morphology of the PFC during early childhood (Huttenlocker, 1979; Huttenlocker & Dabholkar, 1997; Mrzljak, Uylings, van Eden, & Judas, 1990). It is therefore likely that structural and functional changes within this region contribute to the observed age-related improvements in control processes during early childhood.

Middle Childhood into Adolescence

Substantial development of episodic memory has occurred by the beginning of middle childhood, but protracted change continues to take place. In this section, we summarize some of these later developmental changes.

Binding Processes

It was originally thought that binding processes matured fairly early in life, but more recent findings indicate that the HPC, the region known to support binding, exhibits structural change into adolescence (Gogtay et al., 2006). These findings have raised the question of whether additional changes in binding processes occur later in development. A growing body of work suggests that this is the case. For example, Lee and colleagues (2016) assessed the ability to remember three different types of associations: item-item, item-time, and item-space associations in 8- to 11-year-old children and adults. The contribution of age differences in controlled processes, such as encoding or retrieval strategies, was minimized in this paradigm by the use of novel objects with unknown verbal labels and testing over very short delays. As evident in Figure 4.2, developmental differences were found that depended on the nature of the relation, with memory for item-space relations reaching adult levels of performance before memory for item-time relations, which in turn seemed to approximate levels of adult performance before memory for item-item relations.

The fact that these results were observed despite participants engaging in a single

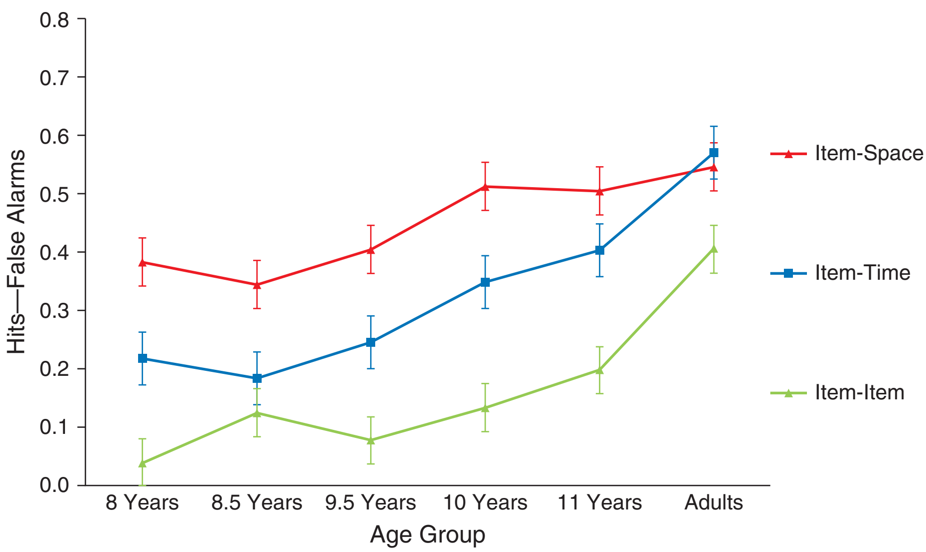


Figure 4.2 This figure presents age-related differences in memory for item-space, item-time, and item-item associations in 8- to 11-year-olds and adults. Error bars represent standard errors. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
Source: Adapted from Lee et al. (2016).

encoding procedure further underscores that binding relations may underlie these differences. It seems that not all binding is created equal, a notion that is supported by other work showing that binding of temporal information appears to follow a more protracted developmental trajectory (see Pathman & St. Jacques, 2014, for review) compared to binding of other types of relations (e.g., Picard, Cousin, Guillery-Girard, Eustache, & Piolino, 2012). For example, Pathman and Ghetti (2014) found developmental differences between 7-year-olds, 10-year-olds, and adults in the extent to which early and obligatory eye movements during retrieval tracked accurate item-time relations. These eye movements are thought to reflect initial binding operations allowing for the reinstatement of a past experience before controlled processes might further assess memory outcomes (Hannula et al., 2010). Pathman and Ghetti also found that individual differences in this eye-movement component predicted overt memory accuracy for temporal order independent from the additional contribution of individual differences in general knowledge of temporal reconstruction processes, again underscoring the role of binding processes in memory improvement into adulthood. Relatedly, these findings open the possibility for a reinterpretation of earlier results showing developmental differences in item-context associations (e.g., item-spatial position) in long-term memory (e.g., Lorsbach & Reimer, 2005) and working memory (e.g., Cowan, Naveh-Benjamin, Kilb, & Saults, 2006). Future work investigating how the brain works to bind different features could therefore provide important insight into the development and basis of episodic memory.

Research showing protracted developmental change in the HPC complements these behavioral findings. This work shows continued myelination within the HPC into adulthood, neurogenesis within the dentate gyrus across the life span (Altman & Das, 1965; Gould & Gross, 2002; Kempermann et al., 2004), and potential increases in entorhinal connectivity with the dentate gyrus into late childhood and adolescence (Abraham et al., 2010). And, neuroimaging work shows an actual relation between hippocampal change (in function and structure) and developmental improvements in episodic memory. For example, in a study testing memory for item-context associations, both 14-year-olds and adults demonstrated hippocampal activation associated with correct retrieval, whereas 8- and 10- to 11-year-olds failed to do so (Ghetti, DeMaster, Yonelinas, & Bunge, 2010). Results from other studies generally align with these findings, showing increased hippocampal specificity for correct retrieval with age (DeMaster & Ghetti, 2013; Demaster, Pathman, & Ghetti, 2013). However, some studies have failed to observe age-related differences in hippocampal function (Güler & Thomas, 2013; Ofen, Chai, Schuil, Whitfield-Gabrieli, & Gabrieli, 2012). Although we cannot be certain as to the source of these discrepancies, it is possible that differences in methodologies play a role (e.g., examining item recognition versus the association between items and some element of the context). We also note recent work suggesting that developmental differences in hippocampal function depend on memory performance level (Sastre et al., 2016). Furthermore, additional findings suggest that age differences in hippocampal activation depend on manipulations that illuminate the nature of developmental changes: Recent comparisons indicated that 8-year-olds fail to recruit the HPC for associative recognition when there is a change between the encoding and retrieval context, but are more likely to recruit this structure compared to adults when there is a match between the encoding and retrieval contexts (DeMaster, Coughlin, & Ghetti, 2016).

Other neuroimaging work suggests that developmental differences in hippocampal function within *specific regions* may contribute to changes in binding processes. In a study by Demaster and Ghetti (2013), anterior hippocampal activation predicted correct retrieval in adults, whereas posterior hippocampal activation predicted correct retrieval in 8- to 11-year-olds. These findings align with work showing developmental differences in hippocampal volume by region. Specifically, Gogtay et al. (2006) have shown reduced anterior hippocampal volume and increased posterior hippocampal volume with age. Interestingly, smaller hippocampal head and tail volume, and larger hippocampal body volume, relate to better memory performance in adults (DeMaster, Pathman, Lee, & Ghetti, 2014). It is therefore possible that developmental differences in binding processes reflect increasing specialization of hippocampal subregions, perhaps due to synaptic pruning in the anterior HPC and neurogenesis in the posterior HPC (Ghetti & Bunge, 2012; Gogtay et al., 2006). Indeed, Lavenex and Banta Lavenex (2013) have suggested that circuits of the human HPC may develop at different rates, leading to differential development of distinct hippocampally-dependent memory processes. Given that children find flexible retrieval especially difficult (Ackerman, 1982; DeMaster et al., 2016; Paz-Alonso, Ghetti, Matlen, Anderson, & Bunge, 2009), it is possible that a late refinement of hippocampal regions supporting flexible binding may contribute to later episodic memory development. This refinement may also contribute to developmental differences that vary by information type (e.g., spatial versus temporal), given evidence of functional segregation within the HPC for processing of spatial versus temporal information (e.g., Ekstrom et al., 2011). Future research should examine the extent to which structural and functional hippocampal development support the behavioral differences discussed in this section.

Control Processes

The continued development of binding processes from middle childhood to adolescence is paralleled by robust increases in control processes. Indeed, a large body of work has provided ample evidence of age-related improvements in children's ability to not only monitor their memories (i.e., how contextually rich and subjectively compelling the mental re-experience of the event is) but also in their utilization of strategies that support the improved regulation of memory encoding and retrieval operations (Bjorklund et al., 2009; Ghetti, 2008; Plude, Nelson, & Scholnick, 1998; Ornstein et al., 2006; Roebers, 2002; Roebers, von der Linden, Schneider, & Howie, 2007; Schneider & Lockl, 2002).

Developmental gains in memory monitoring are thought to emerge from an increased ability to remember past events in a subjectively vivid and compelling manner (a crucial feature of episodic memory; Tulving, 1985; Yonelinas, 1999), as well as from an increased awareness of this experiential feature. Although a basic ability to monitor the strength of one's memories is in place by age 4 (Hembacher & Ghetti, 2014) or 5 (e.g., Ghetti, Qin, & Goodman, 2002; Roebers, Gelhaar, & Schneider, 2004), improvements continue to be observed across childhood. For example, in a study with 7- to 10-year-olds, participants were asked to report whether a set of actions had been enacted, imagined, or never encountered, and to then provide confidence judgments on their decisions (Ghetti, Lyons, Lazzarin, & Cornoldi, 2008). Results revealed age-related improvements in children's ability to calibrate their confidence ratings to predicted differences in memory strength, with 10-year-olds demonstrating greater sensitivity to subtle differences in predicted memory strength compared to 7-year-olds.

These improvements in memory monitoring overlap with improvements in the ability to introspect on the significance of memory states. Older children are not only better at identifying strong versus weak memories, but also at using these judgments to make inferences about the events they represent. Friedman (2007) has demonstrated that sixth graders and adults are aware that the subjective vividness of a remembered event may relate to how long ago the event took place, whereas kindergarteners through fourth graders are not. Ghetti and colleagues have shown similar age-related improvements in children's understanding of memory functioning (Ghetti, Mirandola, Angelini, Cornoldi, & Ciaramelli, 2011). In a study with 6- to 18-year-olds, they asked participants to classify 30 statements as demonstrating either subjective recollection (e.g., "I can tell I saw this picture before because I saw it in green") or subjective familiarity (e.g., "There was definitely a cup in the list, but I can't tell why"). Although all age groups performed well above chance, significant age-related improvements in classification accuracy were observed across the sample. Further, the ability to correctly classify recollection versus familiarity statements was associated with the tendency to be more cautious to claim subjective recollection, suggesting that a better awareness and understanding of subjective memory states reduces individuals' propensity to claim recollection, thereby supporting more selective and accurate memory reports. Work by Koriat and colleagues aligns with this notion (Koriat, Goldsmith, Schneider, & Nakash-Dura, 2001). In their study, children were asked to report details about a story under free-recall and forced-report conditions. Results showed that 7- to 9-year-olds volunteered less accurate information than 10- to 12-year-olds across the free-recall conditions, even when incentivized to report accurate information only. Thus, older children seemed to use the free-recall condition to their advantage, selectively reporting their most accurate information. Together, these and other findings demonstrate developmental improvements in strategic retrieval across childhood and combine with other work (e.g., Sprondel, Kipp, & Mecklinger, 2012) to show gradual age-related change into young adulthood.

Thus far, only control processes that depend on the ability to introspect on one's subjective memory state have been discussed. Other types of control processes, including strategies, also contribute to developmental improvements in episodic memory. For example, Schleepen and Jonkman (2014) have shown significant age-related increases in the ability of 6- to 12-year-olds to use semantic grouping strategies to help with retrieval. In their study, 6- to 7-year-olds did not demonstrate semantic grouping strategies, 8- to 9-year-olds did so only after explicit instruction, and 10- to 12-year-olds did so even without explicit instruction. Similarly, Ghetti and Angelini (2008) have observed age-related improvements in 6- to 10-year-olds' retrieval of line drawings about which they were required to made a semantic judgment during encoding but not in their retrieval of line drawings about which they were required to make a perceptual judgment during encoding. And Daugherty and Ofen (2015) have shown a relation between children's and adults' (ages 8–25 years) belief in the efficacy of shallow versus deep encoding strategies and differences in memory performance with age. In their study, children were more apt to favor shallow encoding strategies whereas adolescents and adults preferred deep encoding strategies. Of importance, developmental differences in the perceived efficacy of deep encoding strategies accounted for better memory improvements with age. Other studies have shown improvements in the selection and utilization of strategies during adolescence. (For review, see Bjorklund et al., 2009.) Together, these findings support additional work showing that increases in the size and organization of one's semantic knowledge occur across childhood and into adolescence, and contribute substantially to the protracted development of episodic memory via control processes (Bjorklund, 1987; Ornstein & Naus, 1985).

The reviewed behavioral findings converge with research showing late age-related changes in the PFC. This region is shown to mature later than other cortical regions (Casey et al., 2005; Huttenlocker, 1990; Kwon, Reiss, & Menon, 2002). At a cellular level, developmental remodeling and changes in the dendritic spine density of the PFC pyramidal neurons have been observed throughout the third decade of life before stabilizing (Petanjek et al., 2011). And, significant age-related increases in prefrontal white matter occur between the ages of 5 and 17 years (Reiss, Abrams, Singer, Ross, & Denckla, 1996). Cross-sectional and longitudinal neuroimaging studies have also shown age-related decreases in the gray matter volume of the PFC starting in early adolescence and extending into the early 20s (Giedd et al., 1999; Mills, Goddings, Clasen, Giedd, & Blakemore, 2014; Suzuki et al., 2005), and activation differences in this region have been associated with developmental improvements in episodic memory for which control processes are likely involved. For example, Paz-Alonso, Ghetti, Donohue, Goodman, and Bunge (2008) asked 8-year-olds, 12-year-olds, and adults to identify words they had studied from a long list of words. Some of the listed words had not been studied but were semantically related to studied words (critical lures). Adults correctly identified more of the studied words than 8-year-olds and recruited the anterior PFC more to distinguish between studied items and critical lures, suggesting that this region contributes to age-related changes in memory monitoring and decision-making processes. In another study, Chiu, Schmithorst, Brown, Holland, and Dunn (2006) examined encoding activation related to subsequent memory across two tasks in 7- to 8-year-olds and 10- to 18-year-olds. One task involved elaborate encoding (i.e., generating verbs related to cue words), whereas the other did not (i.e., passively listening to short stories). Results indicated that prefrontal activation predicted subsequent memory in the verb generation task for both age groups, but only for older children in the story listening task, perhaps due to older children's use of encoding strategies even when not explicitly instructed to do so. Other work by Ofen and colleagues (2007) adds to this developmental picture, showing a positive association between age and DLPFC activation during selective versus passive encoding in 8- to 14-year-olds.

More recent efforts have begun to elucidate prefrontal contributions to the development of memory retrieval operations, including metacognitive operations involved in determining failure to retrieve episodic detail. Using a longitudinal design, Fandakova et al. (2016) examined functional magnetic resonance imaging data from 8- to 9-year-olds, 10- to 12-year-olds, and adults, with two time-points about 1.5 years apart. Participants encoded a series of objects paired with one of three scenes and then performed a retrieval task in which they could elect to report either the target memory detail (i.e., which scene was originally paired with each object) or uncertainty about it. Results indicated that children who engaged the anterior insula more strongly during inaccurate or uncertain responses (compared to accurate memories) exhibited greater longitudinal increases in anterior PFC activation for decisions to report uncertainty (i.e., greater activation for uncertain responses compared to both accurate and inaccurate responses). Interestingly, both of these neural variables predicted improvements in memory accuracy. These results were interpreted as indicating that general neural responses to error might promote the engagement of processes that support metacognitive assessment. More generally, these findings show that effective cognitive control and decision making continue to develop in middle childhood and play an important role in memory development. Altogether, the reviewed work provides strong empirical support for prefrontally-mediated improvements in control processes during middle childhood and beyond.

CONSIDERATIONS FOR FUTURE RESEARCH AND IMPLICATIONS

The research reviewed in this chapter has provided critical insight into the development of episodic memory via binding and control processes. However, there are still many unknowns. This is especially apparent when we try to define the specific features of episodic memory that contribute to difficulties within certain developmental periods. For example, how do limitations in the encoding versus storage versus retrieval of bound representations contribute to normative developmental change? Some of the reviewed work has been used to suggest that limitations with each of these phases contribute to developmental differences across childhood (e.g., Bauer & Leventon, 2013; DeMaster et al., 2016; Pathman & Bauer, 2013), but other work has proposed instead that limitations in storage abilities are particularly consequential in infancy and early childhood (Bauer, 2005), and that developmental differences in retrieval processes might explain later development (e.g., Lloyd et al., 2009). Virtually no study has examined contributions from more than one phase within the same developmental study, and differences in experimental methods make it difficult to make comparisons across studies. Furthermore, the relative contributions of structural and functional changes in the neural substrates of these different phases are uncharted.

We also note that the contributions of binding and control processes to episodic memory development are typically studied in isolation. Although this is understandable—examining both processes within the same study poses unique methodological challenges—it is important that we make continued strides toward examining these processes alongside one another. Not only do both types of processes develop concurrently, but their underlying neural substrates interact with one another in ways that change across development (Lebel & Beaulieu, 2011; Lebel, Walker, Leemans, Phillips, & Beaulieu, 2008; Ofen et al., 2012). Future research should thus attempt to examine the interaction of these processes more closely. Behaviorally, this question has begun to be addressed by examining the effects of manipulations geared toward binding processes on control processes and vice versa (Brehmer, Li, Muller, von Oertzen, & Lindenberger, 2007; Fandakova, Shing, & Lindenberger, 2013a, 2013b). Neuroscience approaches may further contribute via the examination of structural and functional connectivity between hippocampal and cortical regions (Mabbott, Rovet, Noseworthy, Smith, & Rockel, 2009; Wu et al. 2010). It would also be helpful to examine the contributions of regions (e.g., in the posterior parietal cortex) for which hippocampal representations may be recapitulated from further processing by more anterior prefrontal regions (Cabeza et al., 2008; Shimamura, 2010).

More broadly, future research should attempt to integrate extant knowledge of developmental processes with the real-life implications of the protracted trajectory of episodic memory. Critically, it has been posited that the most adaptive function of episodic memory is to support individuals' ability to envision and successfully anticipate future events (i.e., episodic prospection; Suddendorf & Corballis, 2007). From this perspective, the protracted trajectory of episodic memory may be directly related to children's capacity to act in the present in order to prepare for their future.

According to the constructive episodic simulation hypothesis (Schacter & Addis, 2007), the imagination of personal future events is made possible by first accessing details from multiple past memories, and then by recombining those details to construct a mental representation of a novel, plausible, and personal future event. This hypothesis has been supported by research showing that autobiographical memory actually predicts episodic prospection across development (e.g., Busby & Suddendorf, 2005; Coughlin, Lyons, & Ghetti, 2014; Coughlin, Robins, & Ghetti, in press) and that the two abilities are supported by similar neural substrates (Addis, Wong, & Schacter, 2007; Benoit & Schacter, 2015). It is therefore important to consider how the development of binding and control processes might contribute to the development of episodic prospection.

Binding processes are in place early in childhood, but the reviewed literature suggests that these processes exhibit protracted developmental change across childhood and into adolescence (Lorsbach & Reimer, 2005; Riggins, 2014). In addition, binding processes that support flexible retrieval may be particularly late to develop, given that this type of retrieval is especially difficult for children (Gee & Pipe, 1995; Paz-Alonso et al., 2009) and may rely on relatively late developmental change in hippocampal function (DeMaster et al., 2016). Thus, the protracted development of binding processes may contribute to improvements in the ability to flexibly access and recombine multiple memories in the service of episodic prospection. In addition, control processes may help individuals mentally construct a personal future event that is both plausible and coherent (an achievement that is particularly difficult for young children; Coughlin et al., in press). Indeed, it would make sense that the same processes that support individuals' ability to introspect on and regulate their past memories may also contribute to their ability to execute similar mental actions in the service of envisioning their future. To our knowledge, no one has attempted to directly examine the contribution of binding and control processes to the development of episodic prospection, nor to any other high-order cognitive processes thought to be supported by episodic memory. Future research on this topic is important for a more comprehensive characterization of episodic memory development.

The capacity to prospect is only one of the abilities that are supported by the human capacity to remember with vivid detail. It is becoming increasingly clear that the processes supporting episodic memory play important roles in learning across multiple domains, including reading comprehension (Mirandola et al., 2011), inferential reasoning (Zeithamova et al. 2012), knowledge acquisition (Martins et al., 2006), and arithmetic (Menon, 2016). The profound changes in episodic memory are apt to have tangible effects on academic achievement, but little is known about how this occurs and, on the flip side, whether the development of episodic memory can be broadly enhanced, accelerated, or dampened through experiences in the education setting. The time is likely ripe to explore the reciprocal relations between episodic memory development and elementary school education. These investigations are particularly pressing when one considers the staggering number of conditions, ranging from medical diseases to neurodevelopmental disorders, that appear to alter children's memory functioning (Ghetti, Lee, Sims, DeMaster, & Glaser, 2010) and negatively impact their academic achievement potential.

CONCLUSION

The reviewed literature provides an overview of the development of episodic memory, taking into consideration both binding and control processes. We chose to focus on these processes due to their significant contributions to the development of this ability, but note that episodic memory likely emerges from a complex interaction of these and other processes. Research to date has yielded important insight into the development of episodic memory, providing substantial behavioral and neural evidence of its protracted developmental trajectory. The implications of this protracted developmental trajectory are far-reaching, suggesting that many important cognitive abilities, including the ability to envision the future and learn in the classroom, are not fully in place until adolescence or early adulthood. Additional research is therefore needed to further elucidate the neurocognitive development of episodic memory as well as its widespread functional consequences across typical and atypical development.

REFERENCES

Ackerman, B. (1982) Retrieval variability: the inefficient use of retrieval cues by young children. *Journal of Experimental Child Psychology*, *33*, 413–428. doi:10.1016/0022–0965(82)90056-x

Akers, K. G., Restivo, L., Yiu, A. P., De Christofaro, A., Hsiang, H.-L., Wheeler, A. L., . . . Frankland, P. W. (2014). Hippocampal neurogenesis regulates forgetting during adulthood and infancy. *Science*, *344*, 598–602. doi:10.1126/science.1248903

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*, 1363–1377. doi:10.1016/j.neuropsychologia.2006.10.016

Altman, J., & Das, G. D. (1965). Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *Journal of Comparative Neurology*, *124*, 319–335. doi:10.1002/cne.901240303

Baker-Ward, L., Ornstein, P., & Holden D. J. (1984). The expression of memorization in early childhood. *Journal of Experimental Psychology*, *37*, 555–575. doi:10.1016/0022-0965(84)90076-6

Balcomb, F. K., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science*, *11*, 750–760. doi:10.1111/j.1467-7687.2008.00725.x

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development*, *19*, 159–170. doi:10.1016/S0163-6383(96)90015-6

Bauer, P. J. (2005). Developments in declarative memory. *Psychological Science*, *16*, 41–47. doi:10.1111/j.0956-7976.2005.00778.x

Bauer, P. J. (2007). Recall in infancy a neurodevelopmental account. *Current Directions in Psychological Science*, *16*, 142–146. doi:10.1111/j.1467-8721.2007.00492.x

Bauer, P. J., Deboer, T., & Lukowski, A. F. (2007). In the language of multiple memory systems: defining and describing developments in long-term declarative memory. In L. M. Oakes & P. J. Bauer (Eds.), *Short- and Long-term Memory in Infancy and Early Childhood: Taking the First Steps toward Remembering* (pp. 240–270). New York, NY: Oxford University Press.

Bauer, P. J., Doydum, A. O., Pathman, T., Larkina, M., Güler, O. E., & Burch, M. (2012). It's all about location, location, location: children's memory for the "where" of personally experienced events. *Journal of Experimental Child Psychology*, *113*, 510–522. doi:10.1016/j.jecp.2012.06.007

Bauer, P. J., Hertsgaard, L. A., Dropik, P., & Daly, B. P. (1998). When even an arbitrary order becomes important: developments in reliable temporal sequencing of arbitrary ordered events. *Memory*, *6*, 165–198. doi:10.1080/741942074

Bauer, P. J., & Larkina, M. (2014). The onset of childhood amnesia in childhood: a prospective investigation of the course and determinants of forgetting of early-life events. *Memory* (Hove, England), *22*, 907–924. doi:10.1080/09658211.2013.854806

Bauer, P. J., & Leventon, J. S. (2013). Memory for one-time experiences in the second year of life: implications for the status of episodic memory. *Infancy*, *18*, 755–781. doi:10.1111/infa.12005

Bauer, P. J., San Souci, P., & Pathman, T. (2010), Infant memory. *WIREs Cognitive Science*, *1*, 267–277. doi:10.1002/wcs.38

Begus K., Gliga T., & Southgate, V. (2014). Infants learn what they want to learn: responding to infant pointing leads to superior learning. *PLoS ONE*, *9*(10): e108817. doi:10.1371/journal.pone.0108817

Benoit, R. G., & Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia*, *75*, 450–457. doi:10.1016/j.neuropsychologia.2015.06.034

Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: an interpretive review. *Developmental Review*, *7*, 93–130. doi:10.1016/0273-2297(87)90007-4

Bjorklund, D. F., Dukes, C., & Brown, R. D. (2009). The development of memory strategies. In M. L. Courage & N. Cowan (Eds.), *The Development of Memory in Infancy and Childhood* (2nd ed., pp. 145–175). New York, NY: Psychology Press. doi:10.1016/0273-2297(87)90007-4

Brehmer Y., Li, S.C., Müller, V., von Oertzen, T., & Lindenberger, U. (2007). Memory plasticity across the life span: uncovering children's latent potential. *Developmental Psychology*, *43*, 465–478. doi:10.1037/0012-1649.43.2.465

Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development*, *20*, 362–372. doi:10.1016/j.cogdev.2005.05.002

Bushnell, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*, 3–15. doi:10.1038/nrn2459

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. *Nature Reviews Neuroscience*, *9*, 613–625. doi:10.1038/nrn2459

Cabeza, R., & St. Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences*, *11*, 219–227. doi:10.1016/j.tics.2007.02.005

Casey, B. J., Tottenham, N., Liston, C., & Durston, S., 2005. Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*, 104–110. doi:10.1016/j.tics.2005.01.011

Chiu, C. Y. P., Schmithorst, V. J., Brown, R. D., Holland, S. K., & Dunn, S. (2006). Making memories: a cross-sectional investigation of episodic memory encoding in childhood using fMRI. *Developmental Neuropsychology*, *29*, 321–340. doi:10.1207/s15326942dn2902_3

Coughlin, C., Lyons, K. E., & Ghetti, S. (2014). Remembering the past to envision the future in middle childhood: developmental linkages between prospection and episodic memory. *Cognitive Development*, *30*, 96–110. doi:10.1016/j.cogdev.2014.02.001

Coughlin, C., Robins, R. W., & Ghetti, S. (in press). Development of episodic prospection: factors underlying improvements in middle and late childhood. *Child Development*.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, *42*, 1089–1102. doi:10.1037/0012-1649.42.6.1089

Daugherty, A. M., & Ofen, N. (2015) That's a good one! Belief in the efficacy of mnemonic strategies contributes to age-related increase in associative memory. *Journal of Experimental Child Psychology*, *136*, 17–29. doi:10.1016/j.jecp.2015.02.008

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, *208*, 1174–1176. doi:10.1126/science.7375928

DeLoache, J. S., Cassidy, D. J., & Brown, A. L. (1985). Precursors of mnemonic strategies in very young children's memory. *Child Development*, *56*, 125–137. doi:10.2307/1130180

DeMaster, D., Coughlin, C., & Ghetti, S. (2016). Retrieval flexibility and reinstatement in the developing hippocampus. *Hippocampus*, *26*, 492–501. doi:10.1002/hipo.22538

DeMaster, D., & Ghetti, S. (2013) Developmental differences in hippocampal and cortical contributions to episodic retrieval. *Cortex*, *49*, 1482–1493. doi:10.1016/j.cortex.2012.08.004

DeMaster, D., Pathman, T., & Ghetti, S. (2013). Development of memory for spatial context: hippocampal and cortical contributions. *Neuropsychologia*, *51*, 2415–2426. doi:10.1016/j.neuropsychologia.2013.05.026

DeMaster, D., Pathman, T., Lee, J., & Ghetti, S. (2014). Structural development of the hippocampus and episodic memory: developmental differences along the anterior/posterior axis. *Cerebral Cortex*, *24*, 3036–3045. doi:10.1093/cercor/bht160

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 466–503). Cambridge, MA: MIT Press

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Reviews of Neuroscience*, *30*, 123–152. doi:10.1146/annurev.neuro.30.051606.094328

Ekstrom, A. D., Copara, M. S., Isham, E. A., Wang, W. C., & Yonelinas, A. P. (2011). Dissociable networks involved in spatial and temporal order source retrieval. *Neuroimage*, *56*(3), 1803–1813. doi:10.1016/j.neuroimage.2011.02.033

Fandakova, Y., Bunge, S. A., Wendelken, C., Desautels, P., Hunter, L., Lee J. K., & Ghetti, S. (2016). The importance of knowing when you don't remember: neural signaling of retrieval failure predicts memory improvement over time. *Cerebral Cortex*, *23*, 1–13. doi:10.1093/cercor/bhw352

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013a). Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. *Developmental Psychology*, *49*, 1822–1832. doi:10.1037/a0031361

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013b). High-confidence memory errors in old age: the roles of monitoring and binding processes. *Memory*, *21*, 732–750. doi:10.1080/09658211.2012.756038

Fantz, R. L. (1958). Pattern vision in young infants. *Psychological Record*, *8*, 43–48.

Friedman, W. J. (2007). The development of temporal metamemory. *Child Development*, *78*, 1472–1491. doi:10.1111/j.1467–8624.2007.01078.x

Friedman, W. J., & Lyon, T. D. (2005). Development of temporal-reconstructive abilities. *Child Development*, *76*, 1202–1216. doi:10.1111/j.1467-8624.2005.00844.x-i1

Gardiner, J. M., Brandt, K. R., Baddeley, A. D., Vargha-Khadem, F., & Mishkin, M. (2008). Charting the acquisition of semantic knowledge in a case of developmental amnesia. *Neuropsychologia*, *46*, 2865–2868. doi:10.1016/j.neuropsychologia.2008.05.021

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychiatry*, *39*, 3–27. doi:10.1111/1469-7610.00301

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Science*, *11*, 410–419. doi:10.1016/S1364-6613(99)01388-1

Gee, S., & Pipe, M. E. (1995). Helping children to remember: the influence of object cues on children's accounts of a real event. *Developmental Psychology*, *31*, 746. doi:10.1037/0012-1649.31.5.746

Ghetti, S. (2008). Rejection of false events in childhood: A metamemory account. *Current Directions in Psychological Science*, *17*, 16–20. doi:10.1111/j.1467-8721.2008.00540.x

Ghetti, S., & Angelini, L. (2008). The development of recollection and familiarity in childhood and adolescence: evidence from the dual-process signal detection model. *Child Development*, *79*, 339–358. doi:10.1111/j.1467-8624.2007.01129.x

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*, 381–395. doi:10.1016/j.dcn.2012.05.002

Ghetti, S., Castelli, P., & Lyons, K. (2010). Knowing about not remembering: developmental dissociations in lack-of-memory monitoring. *Developmental Science*, *13*, 611–621. doi:10.1111/j.1467-7687.2009.00908.x

Ghetti, S., DeMaster, D., Yonelinas, A., & Bunge, S. (2010). Developmental differences in medial temporal lobe function during memory encoding. *Journal of Neuroscience*, *30*, 9548 9556. doi:10.1523/JNEUROSCI.3500-09.2010

Ghetti, S., Hembacher, E., & Coughlin, C. A. (2013). Feeling uncertain and acting on it during the preschool years: a metacognitive approach. *Child Development Perspectives*, *7*, 160–165. doi:10.1111/cdep.12035

Ghetti, S., Lee, J. K., Sims, C. E., DeMaster, D. M., & Glaser, N. S. (2010). Diabetic ketoacidosis and memory dysfunction in children with type 1 diabetes. *Journal of Pediatrics*, *156*, 109–114. doi:10.1016/j.jpeds.2009.07.054

Ghetti, S., Lyons, K. E., & DeMaster, D.M. (2012). The development of subjective remembering. In S. Ghetti & P. J. Bauer (Eds.), *Origins and Development of Recollection: Perspectives from Psychology and Neuroscience* (pp. 144–167). New York, NY: Oxford University Press.

Ghetti, S., Lyons, K. E., Lazzarin, F., & Cornoldi, C. (2008). The development of metamemory monitoring during retrieval: the case of memory strength and memory absence. *Journal of Experimental Child Psychology*, *99*, 157–181. doi:10.1016/j.jecp.2007.11.001

Ghetti, S., Mirandola, C., Angelini, L., Cornoldi, C., & Ciaramelli, E. (2011). Development of subjective recollection: understanding of and introspection on memory states. *Child Development*, *82*, 1954–1969. doi:10.1111/j.1467-8624.2011.01645.x

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: age, distinctiveness, and subjective experience. *Developmental Psychology*, *38*, 705–718. doi:10.1037/0012-1649.38.5.705

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., . . . Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, *2*, 861–863. doi:10.1038/13158

Gilmore, J. H., Shi, F., Woolson, S. L., Knickmeyer, R. C., Short, S. J., Lin, W., . . . Shen, D. (2012). Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cerebral Cortex*, *22*, 2478–2485. doi:10.1093/cercor/bhr327

Gogtay, N., Nugent, T. F., Herman, D. H., Ordonez, A., Greenstein, D., Hayashi, K. M., . . . Thompson, P. M. (2006). Dynamic mapping of normal human hippocampal development. *Hippocampus*, *16*, 664–672. doi:10.1002/hipo.20193

Gould, E., & Gross, C. G. (2002). Neurogenesis in adult mammals: some progress and some problems. *Journal of Neuroscience*, *22*, 619–623.

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 3492–3496. doi:10.1073/pnas.1515129113

Grossmann, T. (2013). Mapping prefrontal cortex functions in human infancy. *Infancy*, *18*, 303–324. doi:10.1111/infa.12016

Güler, O., & Thomas, K. (2013). Developmental differences in the neural correlates of relational encoding and recall in children: an event-related fMRI study. *Developmental Cognitive Neuroscience*, *3*, 106–116. doi:10.1016/j.dcn.2012.07.001

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 166. doi:10.3389/fnhum.2010.00166

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience*, *27*, 14365–14374. doi:10.1523/JNEUROSCI.4549-07.2007

Hayne, H. (2007). Infant memory development: New questions, new answers. In L. M. Oakes & P. J. Bauer (Eds.), *Short- and Long-Term Memory in Infancy and Early Childhood: Taking the First Steps toward Remembering* (pp. 209–239). New York, NY: Oxford University Press.

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*, 1768–1776. doi:10.1177/0956797614542273

Howe, M. L., & Courage, M. L. (1993). On resolving the enigma of infantile amnesia. *Psychological Bulletin*, *113*, 305–326. doi:10.1037/0033-2909.113.2.305

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain Research*, *163*, 195–205. doi:10.1016/0006-8993(79)90349-4

Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia*, *28*, 517–527. doi:10.1016/0028-3932(90)90031-I

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387*, 167–178. doi:10.1002/(SICI)1096-9861(19971020)387:2%3C167::AID-CNE1%3E3.0.CO;2-Z

Insausti, R., Cebada-Sanchez, S., & Marcos, P. (2010). Postnatal development of the human hippocampal formation. *Advances in Anatomy, Embryology and Cell Biology*, *206*, 1–86. doi:10.1007/978-3-642-03661-3

Josselyn, S. A., & Frankland, P. W. (2012). Infantile amnesia: a neurogenic hypothesis. *Learning and Memory*, *19*, 423–433. doi:10.1101/lm.021311.110

Kempermann, G., Wiskott, L., & Gage, F. H. (2004). Functional significance of adult neurogenesis. *Current Opinion in Neurobiology*, *14*, 186–191. doi:10.1016/j.conb.2004.03.001

Konkel, A., & Cohen, N. J. (2009). Relational memory and the hippocampus: representations and methods. *Frontiers of Human Neuroscience*, *2*, 166–174. doi:10.3389/neuro.01.023.2009

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517. doi:10.1037/0033–295X.103.3.490

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology*, *79*, 405–437. doi:10.1006/jecp.2000.2612

Kreutzer, M. A., Leonard, C., & Flavell, J. H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development*, *40*, 1–60. doi:10.2307/1165955

Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 13336–13341. doi:10.1073/pnas.162486399

Lavenex, P., & Banta Lavenex, P. (2013). Building hippocampal circuits to learn and remember: insights into the development of human memory. *Behavioural Brain Research*, *254*, 8–21. doi:10.1016/j.bbr.2013.02.007

Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *Journal of Neuroscience*, *31*, 10937–10947. doi:10.1523/JNEUROSCI.5302-10.2011

Lebel, C., Walker, L., Leemans, A., Phillips, L., & Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *Neuroimage*, *40*, 1044–1055. doi:10.1016/j.neuroimage.2007.12.053

Lee, J. K., Johnson, E., & Ghetti, S. (2017). Hippocampal development: Structure, function and implications. In D. Hannula & M. Duff (Eds.), *The Hippocampus from Cells to Systems* (pp. 141–166). Dordrecht, the Netherlands: Springer.

Lee, J. K., Nordahl, C. W., Amaral, D. G., Lee, A., Solomon, M., & Ghetti, S. (2015). Assessing hippocampal development and language in early childhood: evidence from a new application of the automatic segmentation adapter tool. *Human Brain Mapping*, *36*, 4483–4496. doi:10.1002/hbm.22931

Lee, J. K., Wendelken, C., Bunge, S. A., & Ghetti, S. (2016). A time and place for everything: developmental differences in the building blocks of episodic memory. *Child Development*, *87*, 194–210. doi:10.1111/cdev.12447

Li, G., Wang, L., Shi, F., Lyall, A. E., Lin, W., Gilmore, J. H., & Shen, D. (2014). Mapping longitudinal development of local cortical gyrification in infants from birth to 2 years of age. *Journal of Neuroscience*, *34*, 4228–4238. doi:10.1523/JNEUROSCI.3976-13.2014

Lloyd, M. E., Doydum, A. O., & Newcombe, N. S. (2009). Memory binding in early childhood: evidence for a retrieval deficit. *Child Development*, *80*, 1321–1328. doi:10.1111/j.1467-8624.2009.01353.x

Lorsbach, T. C., & Reimer, J. F. (2005). Feature binding in children and young adults. *Journal of Genetic Psychology*, *166*, 313–328. doi:10.3200/GNTP.166.3.313-328

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*, 273–293. doi:10.1016/S0028-3932(97)00109-7

Mabbott, D. J., Rovet, J., Noseworthy, M. D., Smith, M. L., & Rockel, C. (2009). The relations between white matter and declarative memory in older children and adolescents. *Brain Research*, *1294*, 80–90. doi:10.1016/j.brainres.2009.07.046

MacFarlane, A. J. (1975). Olfaction in the development of social preferences in the human neonate. *Ciba Foundation, Symposia*, *33*, 103–117. doi:10.1002/9780470720158.ch7

Martins, S., Guillery-Girard, B., Jambaqué, I., Dulac, O., & Eustache, F. (2006). How children suffering severe amnesic syndrome acquire new concepts? *Neuropsychologia*, *44*, 2792–2805. doi:10.1016/j.neuropsychologia.2006.05.022

Menon, V. (2016). A neurodevelopmental perspective on the role of memory systems in children's math learning. In D. B. Berch, D. C. Geary, & K. M. Koepke (Eds.), *Mathematical Cognition and Learning* (pp. 79–107). San Diego, CA: Academic Press. doi.org/10.1016/B978-0-12-801871-2.00004-6

Miller, P. H., & Seier, W. L. (1994). Strategy utilization deficiencies in children: when, where, and why. In H. W. Reese (Ed.), *Advances in Child Development and Behavior* (Vol. 25, pp. 107–156). San Diego, CA: Academic Press. doi:10.1016/S0065-2407(08)60051-8

Mills, K. L., Goddings, A., Clasen, L. S., Giedd, J. N., & Blakemore, S. (2014). The developmental mismatch in structural brain maturation during adolescence. *Developmental Neuroscience*, *36*, 147–160. doi:10.1159/000362328

Mirandola, C., Del Prete, F., Ghetti, S., & Cornoldi, C. (2011). Recollection but not familiarity differentiates memory for text in students with and without learning difficulties. *Learning and Individual Differences*, *21*, 206–209. doi:10.1016/j.lindif.2010.12.001

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, *4*, 257–267. doi:10.1162/jocn.1992.4.3.257

Mullally, S. L., & Maguire, E. A. (2014). Learning to remember: the early ontogeny of episodic memory. *Developmental Cognitive Neuroscience*, *9*(100), 12–29. doi:10.1016/j.dcn.2013.12.006

Murty, V. P., Calabro, F., & Luna, B. (2016). The role of experience in adolescent cognitive development: integration of executive, memory, and mesolimbic systems. *Neuroscience and Behavioral Reviews*, *70*, 46–58. doi:0.1016/j.neubiorev.2016.07.034

Mrzljak, L., Uylings, H. B., van Eden, C. G., & Judas, M. (1990). Neuronal development in human prefrontal cortex in prenatal and postnatal stages. *Progress in Brain Research*, *85*, 185–222. doi:10.1016/S0079-6123(08)62681-3

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinions in Neurobiology*, *7*, 217–227. doi:10.1016/S0959-4388(97)80010-4

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychological Science*, *4*, 7–14. doi:10.1111/j.1467-9280.1993.tb00548.x

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological Review*, *111*, 486–511. doi:10.1037/0033-295X.111.2.486

Newcombe, N. S., Balcomb, F., Ferrara, K., Hansen, M. & Koski, J. (2014), Two rooms, two representations? Episodic-like memory in toddlers and preschoolers. *Developmental Science*, *17*, 743–756. doi:10.1111/desc.12162

Newman, L. S. (1990). Intentional and unintentional memory in young children: Remembering vs. playing. *Journal of Experimental Child Psychology*, *50*, 243–258. doi:10.1016/0022-0965(90)90041-6

Ofen, N., Chai, X. J., Schuil, K. D., Whitfield-Gabrieli, S., & Gabrieli, J. D. (2012). The development of brain systems associated with successful memory retrieval of scenes. *The Journal of Neuroscience*, *32*, 10012–10020. doi:10.1523/JNEUROSCI.1082-11.2012

Ofen, N., Kao, Y. C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., & Gabrieli, J. D. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience*, *10*, 1198–1205. doi:10.1038/nn1950

Ornstein, P. A., Haden, C. A., & Elischberger, H. B. (2006). Children's memory development: Remembering the past and preparing for the future. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan Cognition: Mechanisms of Change* (pp. 143–161). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195169539.003.0010

Ornstein, P. A., & Naus, M. J. (1985). Effects of the knowledge base on children's memory strategies. *Advances in Child Development and Behavior*, *19*, 113–148. doi:10.1016/S0065-2407(08)60390-0

Owens, R. E. (1984). *Language Development: An Introduction*. Columbus, OH: Merrill.

Pathman, T., & Bauer, P. J. (2013). Beyond initial encoding: Measures of the post-encoding status of memory traces predict long-term recall in infancy. *Journal of Experimental Child Psychology*, *114*, 321–338. doi:10.1016/j.jecp.2012.10.004

Pathman, T., & Ghetti, S. (2014). The eyes know time: A novel paradigm to reveal the development of temporal memory. *Child Development*, *85*, 792–807. doi:10.1111/cdev.12152

Pathman, T., & St. Jacques, P. L. (2014). Locating events in personal time: Time in autobiography. In P. J. Bauer & R. Fivush (Eds.), *The Wiley-Blackwell Handbook on the Development of Children's Memory* (pp. 408–426). West Sussex, UK: Wiley-Blackwell.

Paz-Alonso, P., Ghetti, S., Donohue, S. E., Goodman, G. S., & Bunge, S. A. (2008). Neurodevelopmental correlates of true and false recognition. *Cerebral Cortex*, *18*, 2208–2216. doi:0.1093/cercor/bhm246

Paz-Alonso, P., Ghetti, S., Matlen, B. J., Anderson, M. C., & Bunge, S. A. (2009). Memory suppression is an active process that improves over childhood. *Frontiers in Human Neuroscience*, *3*, 24. doi:10.3389/neuro.09.024.2009

Perner, J., Kloo, D., & Stottinger, E. (2007). Introspection and remembering. *Synthese*, *159*, 253–270. doi:10.1007/s11229-007-9207-4

Petanjek, Z., Judaš, M., Šimic, G., Rašin, M. R., Uylings, H. B. M., Rakic, P., & Kostović, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 13281–13286. doi:10.1073/pnas.1105108108

Peterson, C. (2002). Children's long-term memory for autobiographical events. *Developmental Review*, *22*, 370–402. doi:10.1016/S0273-2297(02)00007-2

Picard, L., Cousin, S., Guillery-Girard, B., Eustache, F., & Piolino, P. (2012). How do the different components of episodic memory develop? Role of executive functions and feature-binding abilities. *Child Development*, *83*, 1037–1050. doi:10.1111/j.1467-8624.2012.01736.x

Plude, D. J., Nelson, T. O., & Scholnick, E. K. (1998). Analytical research on developmental aspects of metamemory. *European Journal of Psychology of Education*, *13*, 29–42. doi:10.1007/BF03172811

Ranganath, C. (2010). Binding items and contexts: The cognitive neuroscience of episodic memory. *Current Directions in Psychological Science*, *19*, 131–137. doi:10.1177/0963721410368805

Reiss, A. L., Abrams, M. T., Singer, H. S., Ross, J. L., & Denckla, M.B. (1996). Brain development, gender and IQ in children. A volumetric imaging study. *Brain*, *119*, 1763–74. doi:10.1093/brain/119.5.1763

Richmond, J., & Nelson, C. A. (2007). Accounting for change in declarative memory: A cognitive neuroscience perspective. *Developmental Review*, *27*, 349–373. doi:10.1016/j.dr.2007.04.002

Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: evidence from eye tracking. *Developmental Science*, *12*, 549–556. doi:10.1111/j.1467-7687.2009.00795.x

Richmond, J. L., & Power, J. (2014). Age-related differences in memory expression during infancy: Using eye-tracking to measure relational memory in 6- and 12-month-olds. *Developmental Psychobiology*, *56*, 1341–1351. doi:10.1002/dev.21213

Riggins, T. (2014). Longitudinal investigation of source memory reveals different developmental trajectories for item memory and binding. *Developmental Psychology*, *50*, 449. doi:10.1037/a0033622

Riggins, T., Blankenship, S. L., Mulligan, E., Rice, K., & Redcay, E. (2015). Developmental differences in relations between episodic memory and hippocampal subregion volume during early childhood. *Child Development*, *86*, 1710–1718. doi:10.1111/cdev.12445

Riggins, T., Geng, F., Blankenship, S. L., & Redcay, E. (2016). Hippocampal functional connectivity and episodic memory in early childhood. *Developmental Cognitive Neuroscience*, *19*, 58–69. doi:10.1016/j.dcn.2016.02.002

Roebers, C. M. (2002). Confidence judgments in children's and adult's event recall and suggestibility. *Developmental Psychology*, *38*, 1052–1067. doi:10.1037/0012-1649.38.6.1052

Roebers, C. M., Gelhaar, T., & Schneider, W. (2004). "It's magic!" The effects of presentation modality on children's event memory, suggestibility, and confidence judgments. *Journal of Experimental Child Psychology*, *87*, 320–335. doi:10.1016/j.jecp.2004.01.004

Roebers, C. M., von der Linden, N., Schneider, W., & Howie, P. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, *97*, 117–137. doi:10.1016/j.jecp.2006.12.006

Rovee-Collier, C. (1997). Dissociations in infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review*, *104*, 467–498. doi:10.1037/0033-295X.104.3.467

Rovee-Collier, C., & Gekoski, M. J. (1979). The economics of infancy: A review of conjugate reinforcement. *Advances in Child Development and Behavior*, *13*, 195–255. doi: 10.1016/S0065-2407(08)60348-1

Rubin, D. C. (2000). The distribution of early childhood memories. *Memory*, *8*, 265–269. doi:10.1080/096582100406810

Rugg, M. D., Otten, L. J., & Henson, R. N. (2002). The neural basis of episodic memory: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, *357*, 1097–1110. doi: 10.1098/rstb.2002.1102

Sastre, Marcos, I., II, Wendelken, C., Lee, J. K., Bunge, S. A., & Ghetti, S. (2016). Age- and performance-related differences in hippocampal contributions to episodic retrieval. *Developmental Cognitive Neuroscience*, *19*, 42–50. doi:10.1016/j.dcn.2016.01.003

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *362*, 773–786. doi:10.1098/rstb.2007.2087

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, *8*, 657–661. doi:10.1038/nrn2213

Schacter, D. L., & Moscovitch, M. (1984). Infants, amnesics, and dissociable memory systems. In M. Moscovitch (Ed.), *Infant Memory: Its Relation to Normal and Pathological Memory in Humans and Other Animals* (pp. 173–216). New York, NY: Springer.

Schleepen, T., & M. Jonkman, L. (2014). A longitudinal study of semantic grouping strategy use in 6–11-year-old children: Investigating developmental phases, the role of working memory, and strategy transfer. *Journal of Genetic Psychology*, *175*, 451–471. doi:10.1080/00221325.2014.958126

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2016). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, *29*, 37–51. doi:10.1162/jocn_a_01028

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1–8. doi:10.1016/j.cobeha.2014.07.005

Schneider, W., & Lockl, K. (2002). The development of metacognitive knowledge in children and adolescents. In T. J. Perfect, & B. L. Schwartz (Eds.), *Applied Metacognition* (pp. 224–258). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511489976.011

Seress, L., Ábrahám, H., Tornóczky, T., & Kosztolányi, G. (2001). Cell formation in the human hippocampal formation from mid-gestation to the late postnatal period. *Neuroscience*, *105*, 831–843. doi:10.1016/s0306-4522(01)00156-7

Seress, L., & Ribak, C. E. (1995). Postnatal development and synaptic connections of hilar mossy cells in the hippocampal dentate gyrus of rhesus monkeys. *Journal of Comparative Neurology*, *355*, 93–110. doi:10.1002/cne.903550111

Shimamura, A. P. (2010). Hierarchical relational binding in the medial temporal lobe: The strong get stronger. *Hippocampus*, *20*, 1206–1216. doi:10.1002/hipo.20856

Shing, Y. L., Werkle-Bergner, M., Li, S.-C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A lifespan dissociation. *Journal of Experimental Psychology: General*, *137*, 495–513. doi:10.1037/0096-3445.137.3.495

Sluzenski, J., Newcombe, N., & Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 89–100. doi:10.1037/0278-7393.32.1.89

Sprondel, V., Kipp, K. H., & Mecklinger, A. (2012). Electrophysiological evidence for late maturation of strategic episodic retrieval processes. *Developmental Science*, *15*, 330–344. doi:10.1111/j.1467-7687.2011.01130.x

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, *4*, 232–243. doi:10.1162/jocn.1992.4.3.232

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, *30*, 299–313. doi:10.1017/S0140525X07001975

Suzuki, M., Hagino, H., Nohara, S., Zhou, S., Kawasaki, Y., Takahashi, T., . . . Kurachi, M. (2005). Male-specific volume expansion of the human hippocampus during adolescence. *Cerebral Cortex*, *15*, 187–193. doi:10.1162/jocn.1992.4.3.232

Tsujimoto, S. (2008). The prefrontal cortex: Functional neural development during early childhood. *Neuroscientist*, *14*, 345–358. doi:10.1177/1073858408316002

Tulving, E. (1984). Precis of elements of episodic memory. *Behavioral and Brain Sciences*, *7*, 223–268. doi:10.1017/S0140525X0004440X

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385–398. doi:10.1037/0003-066X.40.4.385

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *The Missing Link in Cognition: Origins of Self-reflective Consciousness* (pp. 3–56). New York, NY: Oxford University Press.

Uematsu, A., Matsui, M., Tanaka, C., Takahashi, T., Noguchi, K., Suzuki, M., & Nishijo, H. (2012). Developmental trajectories of amygdala and hippocampus from infancy to early adulthood in healthy individuals. *PLoS ONE*, *7*(10), e46970. doi:10.1371/journal.prone.0046970

Wellman, H. M. (1977). Preschoolers' understanding of memory-relevant variables. *Child Development*, *48*, 1720–1723. doi:10.2307/1128544

Wellman, H. M. (1988). First steps in the child's theorizing about the mind. In J. Astington, P. Harris, & D. Olson (Eds.), *Developing Theories of Mind* (pp. 64–92). Cambridge, UK: Cambridge University Press.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *12*, 331. doi:10.1037/0033-2909.121.3.331

Wiebe, S. A., Lukowski, A. F., & Bauer, P. J. (2010). Sequence imitation and reaching measures of executive control: A longitudinal examination in the second year of life. *Developmental Neuropsychology*, *35*, 522–538. doi:10.1080/87565641.2010.494751

Wixted, J. T., & Squire, L. R. (2011). The medial temporal lobe and the attributes of memory. *Trends in Cognitive Sciences*, *15*, 210–217. doi:10.1016/j.tics.2011.03.005

Wu, T. C., Wilde, E. A., Bigler, E. D., Yallampalli, R., McCauley, S. R., Troyanskaya, M., . . . Levin, H. S. (2010). Evaluating the relationship between memory functioning and cingulum bundles in acute mild traumatic brain injury using diffusion tensor imaging. *Journal of Neurotrauma*, *27*, 303–307. http://doi.org/10.1089/neu.2009.1110

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1415–1434. doi:10.1037/0278–7393.25.6.1415

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in Human Neuroscience*, *6*, 70. doi:10.3389/fnhum.2012.00070

CHAPTER 5

Development of Cognitive Control Across Childhood and Adolescence

KIKI ZANOLIE AND EVELINE A. CRONE

INTRODUCTION

One of the most salient changes during school-age development is a pronounced increase in the ability to control thought and actions for the purpose of obtaining future goals, also referred to as cognitive control. School-age development encompasses the period of late childhood (6–10 years) and adolescence (10–20 years), and adolescence is defined as a transitional phase between childhood and adulthood. During adolescence, there are significant changes in physical, cognitive, emotional, and social behavior (Blakemore, 2008). The onset of adolescence is marked by the start of puberty, during which time hormone levels rise, triggering a cascade of physical and socioemotional changes, preparing adolescents for independence (Shirtcliff, Dahl, & Pollak, 2009). Puberty starts approximately at the age of 10 to 12 years (on average 1.5 years earlier for girls than for boys), but there are large individual differences between children with respect to when they enter puberty. That is to say, some children may already start puberty at 8 years of age, whereas others do not start puberty until 13 years of age. The end of adolescence is less clearly defined but is generally thought to be the time when individuals obtain mature social goals, which generally occurs around the age of 20 to 25 years (Crone & Dahl, 2012).

Of all cognitive processes, cognitive control is probably the latest to reach adult performance levels, with improvements that are observed over the whole period of child and adolescent development (Diamond, 2013; Huizinga, Dolan, & van der Molen, 2006). One reason for this protracted development is most likely because cognitive control relies on the combination of many lower- and higher-level functions working together. That is to say, the term "cognitive control" refers to a set of cognitive abilities that enable one to control and regulate one's behavior adaptively to meet current and future goals. Cognitive control is comprised of many different components, such as working memory, inhibition, and performance monitoring (Huizinga et al., 2006; Miyake et al., 2000). It is important for school-based functions, such as reading and math, but also for the ability to control impulses and for social interactions. All these functions have shown development improvements across childhood and adolescence.

In the last two decades, much new insight has been gained with respect to understanding the development of cognitive control by relating it to changes in the structure and function of the brain. Longitudinal research on changes in brain structure over development within individuals has shown that regions within the frontal, temporal, and parietal cortices show maturational changes much longer than previously thought, with massive changes occurring in gray matter volume and white matter connections until the early 20s (Gogtay et al., 2004; Mills, Lalonde, Clasen, Giedd, & Blakemore, 2014). Moreover, these studies have shown that different brain structures develop at different rates. For example, several studies have demonstrated that gray matter in cortical areas, reflecting neuronal density and the numbers of connections between neurons, follows an inverted U-shape over development, declining at different ages depending on the region. For this reason, gray matter loss is considered an index of the time course of maturation per region (Casey, Tottenham, Liston, & Durston, 2005). Intriguingly, the most protracted development is observed in the prefrontal cortex (PFC) and parietal cortex, regions of the brain that are consistently implicated in cognitive control (Casey, 2015).

Although it is still debated what the underlying mechanisms associated with a reduction in gray matter volume in cortical areas are (see Paus, Keshavan, & Giedd, 2008), it is thought that these reductions reflect synaptic reorganization and/or increases in white matter integrity (Paus et al., 2008). This idea that cortices continue to undergo synaptic pruning across adolescence is supported by histological studies of the postmortem human brain (Huttenlocher & Dabholkar, 1997; Petanjek et al., 2011).

More direct ways to investigate the relation between brain development and cognitive control is through examining brain responses while individuals perform cognitive tasks. Two methods have been most useful to examine these relations: functional magnetic resonance imaging (fMRI), a method with good spatial resolution (Huettel, Song, & McCarthy, 2004), and event-related potentials (ERPs) derived from the electroencephalogram (EEG), a method with good temporal resolution (Segalowitz, Santesso, & Jetha, 2010a).

In this chapter, we discuss new insights on functional development of brain regions supporting key aspects of cognitive control and show that this functional development has consequences not only for academic achievement but also for affective and social development, specifically social decision making, which requires inhibiting selfish impulses for the purpose of fairness and reciprocity.

We describe the development of three key aspects of cognitive control: working memory, inhibition, and performance monitoring. Whereas working memory and inhibition are important for keeping information in and out of mind, performance monitoring plays a central role in the development of behavioral adjustment. Performance monitoring is therefore particularly important for rapid adaptation to a variety of changing social environments, such as forming new friendships, changing schools, and making future-oriented choices.

DEVELOPMENTAL CHANGES IN COGNITIVE CONTROL AND THE ROLE OF THE PREFRONTAL CORTEX

Cognitive Control and the Prefrontal Cortex

It has been well conceptualized for over a century that the PFC plays an important role in cognitive control (Milner, 1963). Early studies of patients with PFC lesions showed that these patients have difficulties with cognitive flexibility. This was demonstrated using the now-famous Wisconsin Card Sorting Task (Milner, 1963). In this task, individuals are asked to sort decks of cards using a certain criterion that they need to discover themselves, such as color, shape, or number. Each sort is followed by positive or negative feedback, and as such, the participant needs to pay attention to the outcomes of their behavior and use the feedback to discover rules in the task. After a certain number of correct sorts, however, the sorting rule changes (e.g., unbeknownst to the participants sorting to color, this sort no longer results in positive feedback, but the correct sorting dimension is now shape). Participants need to respond flexibly to this change and need to update behavior accordingly. It was found that especially patients with damage to the dorsolateral prefrontal cortex (DLPFC) and medial prefrontal cortex (mPFC) have difficulty with updating their behavior (Barcelo & Knight, 2002). Interestingly, young children show similar deficits on this task (Crone, Ridderinkhof, Worm, Somsen, & van der Molen, 2004), and over the course of adolescence, children become more successful in performing the Wisconsin Card Sorting Task (Huizinga et al., 2006). It is clear that this task requires many different control processes, such as working memory (keeping the sorting rule active), response inhibition (inhibit responding according to the old rule), and performance monitoring (paying attention to the feedback).

Many studies in adults have used neuroimaging to unravel how different regions in the PFC are involved in these separable processes involved in performing complex cognitive control tasks. These studies have consistently indicated that both the ventrolateral cortex and the DLPFC, along with the parietal cortex, are active during working memory tasks (see Figure 5.1; D'Esposito & Postle, 2015; Klingberg, 2010). In contrast, response inhibition typically engages the right ventrolateral (VL) PFC and pre-supplementary motor area (pre-SMA; Aron, Robbins, & Poldrack, 2014). Finally, performance monitoring typically engages both lateral and mPFC, although the lateral PFC is more strongly engaged in rule searching based on feedback, whereas the mPFC is more strongly engaged in processing rule-violating feedback (Zanolie, Van Leijenhorst, Rombouts, & Crone, 2008a).

Recently, developmental neuroimaging studies have examined how the regions within

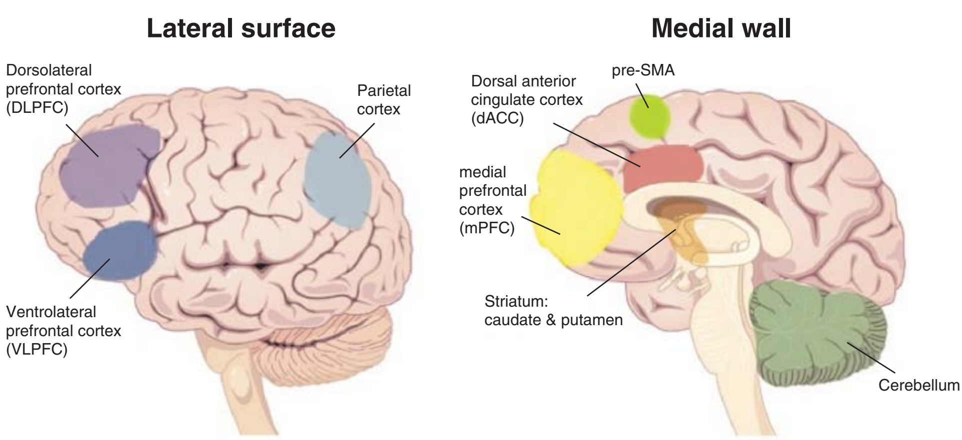


Figure 5.1 Brain regions implicated in cognitive control and social decision making. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174

the PFC are related to the development of cognitive control. Next we describe these studies in the domains of working memory, inhibition, and performance monitoring.

Development of Working Memory

One of the most studied components of cognitive control is working memory, which refers to the ability to keep information online while ignoring irrelevant information (Baddeley & Logie, 1999). The ability to keep information in mind is essential for a wide variety of cognitive abilities, including mathematical calculation, reading, problem-solving, and reasoning (Bayliss, Jarrold, Baddeley, & Gunn, 2005; Swanson, 2004). In fact, working memory capacity predicts school performance, such as reading and arithmetic (Hitch, Towse, & Hutton, 2001). Moreover, it has been demonstrated that working memory shows a strong connection to fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999). The ability to keep information in working memory matures slowly during childhood and is thought to be the driving force behind cognitive development (e.g., Casey, Giedd, & Thomas, 2000; Diamond, 2002).

Many studies have used the model of Baddeley and Hitch (1974) as a framework for understanding working memory. According to this model, working memory comprises of a central executive and two slave systems, the phonological loop and the visuospatial sketchpad. The phonological loop is specialized in processing language-based information, whereas the visuospatial sketchpad is specialized in processing visuospatial information. Assessment of these two systems usually is made by means of short-term memory tasks in which small amounts of information are to be held and reproduced, with no additional cognitive demands (e.g. digit span, word recall, Corsi blocks, and visual-patterns tasks). The central executive controls the allocation of resources between the phonological loop and the visuospatial sketchpad and is able to update and manipulate the content of memory when new and relevant information is processed.

Assessment of the central executive typically involves experimental tasks that require participants to update or manipulate information currently held in working memory. Such tasks include the listening span, counting span, backward digit span, and *n*-back task. In fMRI paradigms, the *n*-back task is frequently used to investigate working memory capacity. (For a meta-analysis, see Owen, McMillan, Laird, & Bullmore, 2005.) In the *n*-back task, participants are presented with a series of stimuli (e.g., letters or words). Participants are required to maintain these stimuli online and compare it with *n* stimuli back. For example, in the 3-back letter version, subjects are required to compare each presented letter with the letter presented 3 before. Participants need to indicate whether the letters are the same or different. In order to perform well on this task, participants need to maintain information in memory and, at the same time, manipulate and update its content.

In adult brain imaging studies, it has been shown that maintenance and manipulation of information in working memory are associated with activation in the ventral and dorsal parts of the PFC. Typically, VLPFC is involved in maintenance of information in working memory, while the DLPFC also is recruited when manipulation of information is needed (Crone, Wendelken, Donahue, van Leijenhorst, & Bunge, 2006; D'Esposito, Postle, Ballard & Lease, 1999; Owen, Evans, & Petrides, 1996; Sakai & Passingham, 2002; Smith & Jonides, 1999; Wagner, Maril, Bjork, & Schacter, 2001). Besides the DLPFC, the superior parietal cortex also is implicated in tasks involving manipulation. (For a meta-analysis, see Wager & Smith, 2003.)

To test the neurocognitive development of working memory, studies have compared neural activity in children, adolescents, and adults when they performed a working memory task in the MRI scanner. One way of measuring working memory is by varying the amount of time between stimulus presentations and retrieving information, or by varying the number of items that have to be held in working memory. This type of working memory is referred to as working memory maintenance. Several researchers have shown that activity in the lateral PFC and parietal cortex during working memory maintenance increases from childhood to adulthood (Klingberg, Forssberg, & Westerberg, 2002; O'Hare, Lu, Houston, Bookheimer, & Sowell, 2008). Others have shown that, in children, activity is likely to be diffuse across different brain regions, whereas in adolescence, it tends to be more restricted to the specific neural regions that show task-relevant activation as in adults (Ciesielski, Lesnik, Savoy, Grant, & Ahlfors, 2006; Geier, Garver, Terwilliger, & Luna, 2009; Libertus, Brannon, & Pelphrey, 2009; Scherf, Sweeney, & Luna, 2006).

Age differences become even larger when participants are asked to reorganize the information in working memory. This type of working memory is measured by the so-called manipulation task. One brain imaging study compared three age groups—children ages 8 to 12 years, adolescents ages 13 to 17 years, and young adults—performing a working memory task in which three pictures were shown and needed to be remembered (Crone et al., 2006). In the forward condition, the participants were required to remember the pictures in the presented order during a short delay period. (See Figure 5.2.) In order to perform this condition well, participants only needed to maintain the presented information in working memory. However, in the backward condition, participants were required to remember the pictures in reverse order; therefore, they had to manipulate *and* maintain the information during the delay period. Maintenance of information in working memory (forward condition) was associated with increased VLPFC activation showing the same pattern for adults, adolescents,

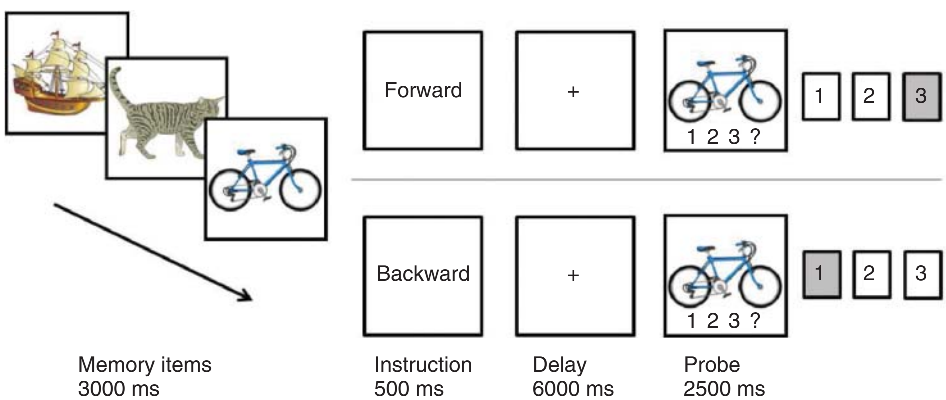


Figure 5.2 Working memory task in which participants were instructed to remember items in either forward or backward order. Correct answers are shaded. This task measures maintenance and manipulation abilities of working memory. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174
SOURCE: Adapted from Crone et al. (2006).

and children. However, in the backward condition, when information needed to be manipulated, adults and adolescents recruited the DLPFC in addition to the VLPFC, but the 8- to 12-year-old children did not. It is important to note that increased brain activity does not reflect a stronger activation overall but a stronger brain activation for a specific contrast. In this study, for instance, activation during manipulating information was compared to maintaining information (backward condition > forward condition).

The number of correct responses in the backward condition correlated with the activation in the DLPFC, strengthening the hypothesis that the DLPFC is important for improvements in performance on working memory manipulation (Crone et al., 2006). In a subsequent study, it was found that this effect was independent of the number of items that had to be held within working memory, suggesting that the effects are specific for manipulation and not related to task difficulty (Jolles, Kleibeuker, Rombouts, & Crone, 2011). Similar results were obtained when researchers presented this task in a visuospatial *n*-back version. In this task, it was found that both activity in the lateral PFC and the parietal cortex increased linearly across ages 7 to 22 years (Kwon, Reiss, & Menon, 2002; Spencer-Smith et al., 2013), possibly related to adults being better able to activate these regions over a sustained period of time (Brahmbhatt, White, & Barch, 2010).

Together these studies provide evidence that the development of separable cognitive control functions within working memory are associated with differential maturational patterns of subregions within the lateral PFC. The differential engagement of subregions of the lateral PFC is in accordance with structural changes across development within the brain. Within the PFC, gray matter reduction is observed earlier for the VLPFC compared to the DLPFC, such that the VLPFC reaches mature levels at adolescence, whereas the DLPFC continues to have gray matter loss until young adulthood (Gogtay et al., 2004). More important, a study of 951 participants between the ages of 8 and 22 years showed that activity in the lateral PFC mediates the relation between age and working memory (Satterthwaite et al., 2013), suggesting that children have more difficulty with working memory due to the protracted development of the PFC.

Apparent contradictory findings with regard to developmental brain activity patterns while maintaining information online might be due to task-specific processing requirements, such as visuospatial information processing (Klingberg et al., 2002) versus verbal information processing (Crone et al., 2006; O'Hare et al., 2008). Also, working memory load may account for differences in brain activity patterns (O'Hare et al., 2008).These task-specific processes and differences in working memory load also may contribute to differential developmental patterns in behavior. Although overall behavioral findings suggest that children and adolescents are capable of performing at adult levels, however, their ability to stay focused on the task and to monitor their behavior still improves until adulthood (Luna, Padmanabhan, & O'Hearn, 2010).

Development of Response Inhibition

Inhibition has been abundantly researched in development psychology. Researchers often use computer tasks to measure very precisely how children and adults succeed in stopping their responses. In the go/no-go task (for a review, see Hester, Fassbender, & Garavan, 2004), participants are asked to press a button if a certain letter is shown, for example, the letter *X*. This letter is shown in quick succession, but participants are instructed that when the letter *Y* is presented, the button may not be pressed. This is especially hard for preschoolers (Diamond, 2013), but children up to the age of 12 or so also have more trouble inhibiting their behavior than adults do (Casey, 2015; Schel & Crone, 2013).

Another way of measuring inhibition is with the stop-signal task (Rubia, Smith, Taylor, & Brammer, 2007). In this task, participants are instructed to respond to green arrows pointing left or right by pressing a left or right button with their index fingers. But when the arrow turns red, participants are not allowed to respond. This is relatively easy when the arrow turns red right away, but it is more difficult when the arrow remains green for a while first and turns red only when participants are about to press the button. By varying the time between showing the arrow and turning the arrow red, researchers can determine how much time someone needs in order to be able to stop successfully. This time is called the stop signal reaction time. Between the ages of 3 and 6, the stop-signal reaction time is slow, meaning that they need more time compared to adults to be able to inhibit their initial response (Lee, Lo, Li, Sung, & Juan, 2015). This reaction time does not reach adult levels until the ages of 12 to 14 at least (van den Wildenberg & van der Molen, 2004). Until that age, children and adolescents have more difficulty stopping than adults (Schel, Scheres, & Crone, 2014).

Damage to the right ventral area of the lateral PFC leads to great difficulty with stopping (Aron et al., 2004). Healthy adults show activity in this area when performing a go/no-go task or a stop-signal task (Aron et al., 2004, 2014).

Researchers have investigated the development of this inhibition area between the ages of 8 and 12 and the ages of 18 and 25 using a go/no-go task or stop-signal tasks. The most important finding was that when 8- to-12-year-olds perform a go/no-go task during an fMRI scan, they show less activity in the right ventral part of the lateral PFC than do 18- to 25-year-olds (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Durston et al., 2006; Rubia et al., 2006; Tamm, Menon, & Reiss, 2002). Interestingly, the younger children often show additional activity in a different area of the PFC, such as the dorsal area of the PFC (Booth et al., 2003; Durston et al., 2006). Similar results have been obtained for the stop-signal task (Rubia et al., 2007, 2013; Vink et al., 2014). These findings suggest again a shift from diffuse to localized activity.

Development of Performance Monitoring

Performance monitoring is of pivotal importance to rapidly adapt to a changing (social) environment. It is important not only to be able to monitor performance in order to adjust behavior after committing an error but also when making future-oriented choices and forming new friendships. A wide variety of tasks have been used to study the neural processes involved in the development of performance monitoring, ranging from cognitive learning tasks (e.g., Hajcak, Moser, Holroyd, & Simons, 2006; Holroyd & Coles, 2002) to socioaffective feedback tasks (e.g. Fareri & Delgado, 2014; Somerville, Heatherton, & Kelley, 2006). In performance monitoring studies, participants typically have to make a single or binary choice when presented with a stimulus. Upon this choice, performance feedback is given, which can have a positive valence when the choice is correct or a negative valence when the choice is incorrect. Studies in adults have shown great involvement of the dorsomedial PFC, more specifically the pre-SMA and the dorsal anterior cingulate cortex (dACC) (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004).

Performance monitoring can be subdivided into internal and external feedback processing or monitoring. The event-related potential (ERP) technique has proven to be very useful in investigating these types of monitoring, especially internal monitoring. ERP studies have shown that internal feedback processing, when errors are processed, is associated with a negative ERP component, the error-related negativity (ERN), which peaks approximately between 50–100 ms after an error is committed. This component is visible even before external feedback may be presented. Traditionally the ERN is examined by means of speeded response tasks in which conflicting stimulus-response mappings are likely to occur equally, such as the go/no-go and Flanker tasks. In the go/no-go task, prepotent responses to target stimuli need to be inhibited, while in the Flanker task, stimuli need to be discriminated with congruent or incongruent flanking stimuli. The source of the ERN is estimated in the dACC by using source localization methods (Carter & van Veen, 2007; Holroyd & Coles, 2002).

Compared to the ERN, which reflects the activation of an internal feedback processing or monitoring system, feedback-related negativity (FRN) reflects activity associated with external monitoring (Gentsch, Ullsperger, & Ullsperger, 2009). External feedback processing is associated with a negative ERP component, the FRN (also referred to as the medial frontal negativity, MFN). This component peaks approximately 250–350 ms after feedback is presented. The source of the FRN has been estimated in both the dACC (van Noordt & Segalowitz, 2012) and in the basal ganglia (Foti, Weinberg, Dien, & Hajcak, 2011). These sources are consistent with the hypothesis that this component reflects activity of phasic increases and decreases in midbrain dopamine systems, which signal to the dACC that outcomes are better or worse than expected (Holroyd & Coles, 2002). However, there is an ongoing debate with respect to whether the FRN reflects a response to negative feedback (Segalowitz, Santesso, Murphy, Homan, & Chantziantoniou, 2010b), reflects a reward-related positivity that is absent on negative feedback (i.e., nonreward) trials (Foti et al., 2011), or is a response to saliency of feedback irrespective of valence (Santesso, Dzyundzyak, & Segalowitz, 2011).

In fMRI paradigms, it was previously found that processing negative feedback compared to positive feedback is associated with an increase in pre-SMA and dACC activity (Holroyd et al., 2004; Zanolie, et al., 2008b), where neural responses are larger when negative feedback is unexpected (Zanolie et al., 2008a). Typically, activity in the pre-SMA and dACC during performance monitoring tasks is accompanied by activation in lateral parts of the PFC, which together have been interpreted as areas involved in behavioral adjustment (Carter & van Veen, 2007; Kerns et al., 2004). These results suggest that the dorsomedial prefrontal cortex (DMPFC) is involved in cognitive error and negative feedback processing.

During adolescence, children and adolescents improve in the ability to monitor their own actions. Particularly, abilities to monitor, evaluate, and adjust behavior according to changing environmental and social demands improve rapidly. Internal monitoring of actions has been studied to a great extent using ERPs, by focusing on the ERN. Typically internal signals are activated whenever a person notices that an error has been committed (van Noordt & Segalowitz, 2012). Studies that have investigated the ERN in a developmental population have found that the ERN is present, but small, in children between ages 3 and 12 years (Grammer, Carrasco, Gehring, & Morrison, 2014; Richardson, Anderson, Reid, & Fox, 2011;

Torpey, Hajcak, Kim, Kujawa, & Klein, 2012; van Meel, Heslenfeld, Rommelse, Oosterlaan, & Sergeant, 2012). Between late childhood (10–12 years) and early adulthood, the ERN becomes larger after committing an error (Santesso, Segalowitz, & Schmidt, 2006; van Meel et al., 2012). During late adolescence and adulthood, there are continued changes in the manifestation of the ERN (Ladouceur, Dahl, & Carter, 2007; for a review, see Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013).

The development of external feedback monitoring has been studied by measuring the FRN (e.g., Crowley et al., 2009; Holroyd, Baker, Kerns, & Muller, 2008; van Meel, Oosterlaan, Heslenfeld, & Sergeant, 2005; Segalowitz et al., 2010b). Only few studies have directly compared the FRN in healthy, normally developing children with adults (Eppinger, Mock, & Kray, 2009; Santesso et al., 2011); other studies only included an adolescent group of participants (Crowley et al., 2013; Hämmerer, Li, Müller, & Lindenberger, 2011; Zottoli & Grose-Fifer, 2012). Hämmerer et al. (2011) found that the FRN was largest in children and decreased with age. Also, the differences between FRN amplitudes after losses and gains were smallest in the youngest age group. Last, there seem to be latency differences in the FRN between age groups, such that the FRN peaks later in time for 10- to 12- and 13- to 14-year-olds compared to 15- to 17-year-olds (Crowley et al., 2013). Also, Zottoli and Grose-Fifer (2012) found a trend toward latency differences, where the FRN was later in time for 14- to 17-year-olds compared to 22- to 26-year-olds. These results seem to suggest that, during childhood into adolescence, the ability to use external feedback more efficiently is still developing, such that the neural processes underlying the FRN are still developing from childhood to adulthood. A tentative idea is that there might be age-related differences in the efficiency to differentiate between positive and negative feedback related to the developing dopamine system (Zottoli & Grose-Fifer, 2012).

External feedback monitoring has also been studied extensively using brain imaging techniques. Children and early adolescents (8–12 years) show remarkable improvements in behavioral adjustment after receiving cues or feedback signaling the need to change their current behavior. For example, Crone, Zanolie, Van Leijenhorst, Westenberg, and Rombouts (2008) measured the neural activity of three age groups (8–11 years, 14–15 years, and 18–25 years) while performing a rule-learning task, which was based on the principles of the Wisconsin Card Sorting Task. In this rule-learning task, participants were required to sort a stimulus in one of four locations. After each sort, participants received feedback whether the sort was correct (positive feedback) or incorrect (negative feedback). Participants had to use this trial-and-error learning in order to find the correct sorting rule. As soon as the rule was applied for a variable number of trials, the rule changed unexpectedly to a new rule. This change signaled participants that they needed to figure out the new rule by using the given feedback. Developmental comparisons showed that 8- to 11-year-olds activated the ACC/pre-SMA for all types of negative feedback, whereas 14- to 15-year-olds and 18- to 25-year-olds activated this region specifically after unexpected negative feedback signaling a rule shift and, therefore, a need to adjust current behavior. These results show that the pattern of activation changed between early and midadolescence. Developmental increases in neural activity following negative feedback were also observed in the DLPFC; however, this region showed a more protracted development with continued changes between midadolescence and adulthood (van den Bos, Güroĝlu, van den Bulk, Rombouts, & Crone, 2009; van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008).

In a rule-learning paradigm, Peters, Braams, Raijmakers, Koolschijn, and Crone (2014) aimed to pinpoint the exact developmental time point at which the neural response after negative performance feedback reaches adult levels. To this end, they included 268 participants in the age of 8 to 25 years. The researchers found that the developmental pattern in learning from negative feedback, and associated activity in dACC/pre-SMA and DLPFC following negative feedback, increased until age 14, after which it stabilized. Strong correlations between brain responses and behavior on the task showed that this dACC/pre-SMA and DLPFC network was involved in performing the task well, but there was also unique activity associated with age-related changes. These neural patterns of protracted development in dACC/pre-SMA and DLPFC have been observed in a variety of executive control paradigms, such as studies examining response inhibition and working memory, and therefore are to be interpreted as the maturation of a cognitive learning process.

In conclusion, several studies examining performance monitoring based on internal monitoring (ERN) as well as external monitoring (FRN and processing feedback) have consistently shown a developmental growth in abilities. Not only the ability to monitor errors but also the ability to distinguish between different types of feedback develops from childhood through adolescence. More specifically, between childhood and adolescence, the capacity for internal performance monitoring increases. Additionally, adolescents become more successful in filtering the informative value of feedback and, as a consequence, use feedback more efficiently in order to learn. These behavioral improvements are accompanied by functional changes in the brain, such that activity in the DMPFC (more specifically the dACC/pre-SMA) and lateral PFC increases across age. These developmental patterns have been attributed to changes in executive functions and protracted development of the DMPFC and DLPFC.

DEVELOPMENTAL CHANGES IN AFFECTIVE DECISION MAKING

Delay of Gratification

Many of the decisions we make in daily life do not involve only simple deliberations. Many times our decisions are the result of a complex interplay between choosing immediate benefits or long-term outcomes. Specifically this weighing of short- versus long-term consequences of choices seems to undergo pronounced developmental changes in adolescence. For example, in a card-playing task in which children, adolescents, and adults can choose between cards with an immediate high reward but high long-term losses or immediate low reward but small long-term losses, children prefer immediate rewards whereas adults prefer delayed rewards. This choice pattern changes during adolescence, when teenagers are learning to make long-term choices, but even in 16- to-18-year-olds, the choice pattern is not quite as targeted on the long term as it is in adults aged 20 to 25 (Cauffman et al., 2010; Crone & van der Molen, 2004; Hooper, Luciana, Conklin, & Yarger, 2004).

The orbitofrontal cortex in particular plays an important role in controlling responses to reward stimuli (O'Doherty, 2011), as was demonstrated earlier in neuropsychological research, but which was confirmed in white matter tract studies in healthy adults (Peper et al., 2013; van den Bos, Rodriguez, Schweitzer, & McClure, 2014). Whereas function activation patterns inform us about activity in a specific region, white matter tracts inform us about the structural connections between regions. Other areas that play roles in weighing choices are the lateral frontal areas, which are of considerable importance for keeping track of long-term goals (Casey, 2015). For example, when a choice has to be made between an immediate, quick reward or a possible, larger reward on the long term, the emotion-related areas of the brain are active when the quick reward is chosen, whereas the lateral frontal cortex areas are active when the long-term alternatives are chosen (McClure, Laibson, Loewenstein, & Cohen, 2004).

A well-known task that measures these types of decisions in the laboratory is the delay of gratification task (also known as temporal discounting task). This task asks participants to choose between a smaller immediate reward (e.g., 5 euros today) or a larger delayed reward (e.g., 8 euros in 2 weeks). The more impulsive individuals typically choose more for the immediate reward. It is often found that children make more impulsive choices and that the ability to delay gratification (or make long term choices) increases over the course of adolescence (Achterberg, Peper, Van Duijvenvoorde, Mandl, & Crone, 2016; Banich et al., 2013; Steinbeis, Haushofer, Fehr, & Singer, 2016). Interestingly, when adolescents make immediate choices, they show stronger activity in the ventral striatum/nucleus accumbens than adults (Christakou, Brammer, & Rubia, 2011). Overcoming responses to immediate reward is associated with strong functional coupling between the regulating DLPFC and the reward-valuing ventromedial PFC, a connection that becomes stronger with increasing age (Steinbeis et al., 2016).

Moreover, studies that have looked at white matter tracts between the striatum and the frontal cortex have shown that the stronger these connections, the less impulsive people are. White matter tracts become stronger between childhood and adulthood, which explains, at least in part, the developmental changes in delay of gratification (Achterberg et al., 2016; Olson et al., 2009; van den Bos et al., 2015). Thus, the development of the PFC, which is important for several aspects of cognitive control, is most likely also of crucial importance when making patient, long-term beneficial decisions.

DEVELOPMENTAL CHANGES IN COGNITIVE CONTROL AND SOCIAL DECISION MAKING

Development of the Social Brain

Adolescence is seen as a very important transition period for the development of concern for others and social values, which is strongly tied to the development of cognitive and socioaffective abilities. Here we present the argument that cognitive control is also involved in social decision making. Before describing this in more detail, it is important to have a general overview of brain regions that are involved in social reasoning, which include mentalizing about thoughts of others and mentalizing about own benefits.

Mentalizing, the ability to understand the mental state of oneself and another, is closely tied to cognitive and socioaffective development. One of the most important milestones of mentalizing capacities is the development of social perspective-taking abilities (Saxe & Kanwisher, 2003). The term “perspective-taking” refers to the ability to understand intentions, considerations, and goals of others from the point of view of the other person. The core components of perspective-taking mature before a child reaches the age of 5, leading to a “theory of mind” (Wellman, Cross, & Watson, 2001). However, development of these perspective-taking abilities does not stop there. In experimental designs, it was shown that the ability to understand and consider intentions of others in social interactions gradually develops during childhood and adolescence. For example, Dumontheil, Apperly, and Blakemore (2010) had participants view a set of shelves with objects, which they needed to move by instructing another person. However, this other person was not able to see all the objects. Therefore, in order to move the objects correctly, participants needed use the perspective of the other person. These participants showed that the ability to view a certain situation from the point of view of another person continues to develop in adolescence. In adults, perspective taking is associated with activity in the temporoparietal junction, superior temporal sulcus, and DMPFC (Denny, Kober, Wager, & Ochsner, 2012; Van Overwalle, 2009).

The second important component of mentalizing is self-referential processing, which involves comparing consequences for oneself to consequences for others (Rilling & Sanfey, 2011). In adults, self-referential processing is associated with activity in the ventro-mPFC (Amodio & Frith, 2006; Denny et al., 2012). Activation patterns in the ventro-mPFC, temporoparietal junction, superior temporal sulcus, and DMPFC—areas that, together, are referred to as the social brain—change remarkably across adolescence and may influence adolescents' perspective-taking and self-referential abilities in decision making (Blakemore, 2008).

Interestingly, in this perspective, the DLPFC is not only important for cognitive control and affective control but has also been identified as one of the key brain regions associated with social decision making. As such, the DLPFC may contribute to the differential pattern observed in childhood and adolescence regarding social decisions.

Next we provide evidence for a role of cognitive control in one specific aspect of social decision making, the consideration of fairness.

Development of Self–Other Perspectives in Consideration of Fairness

Cognitive control is a very important component of social interactions. Several studies have demonstrated the role of cognitive control by using social dilemma's, or economic games. During social decisions when goods need to be divided between two individuals, two motivational aspects are of importance: interest in one's own benefit and concern for others (Van Dijk & Vermunt, 2000). Over the past decades, many studies have investigated the development of fairness in different types of bargaining contexts. These paradigms stem from social and economic psychology and are particularly valuable in studying the role of cognitive control in considerations of fairness.

Trust Game

In the Trust Game, there are two players with a certain sum of money, the stake (see Figure 5.3; Berg, Dickhaut, & McCabe, 1995). The first player decides independently either to divide the original stake or to trust the second player with the money. If the first player trusts the second player, the stake is tripled. However, the second player now has the power to divide all of the money (the tripled stake) as he or she wishes. The second player can either reciprocate the trust given by the first player by dividing the money relatively fairly between self and the first player. Or the second player can defect and keep the profit, therefore giving nothing or only a small amount of the money back to the first player. Although there are many variations to the game, it usually involves a single transaction with an unknown other to avoid reputation effects.

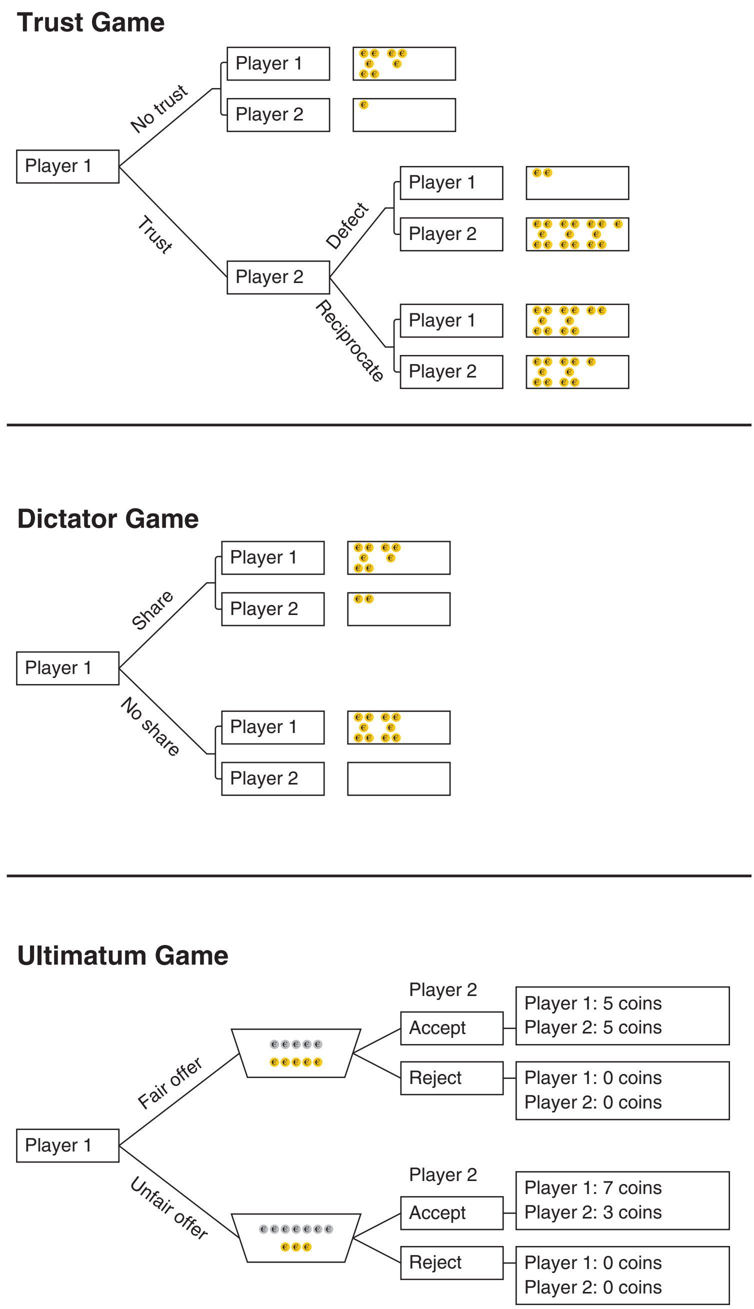


Figure 5.3 Three types of social bargaining games used to study the role of cognitive control in considerations of fairness. In the *Trust Game*, player 1 can decide not to trust and the game ends. If player 1 decides to trust, player 2 ends the game by either defecting or reciprocating. The stakes are represented by a number of coins in boxes next to the names of the players. In the *Dictator Game*, player 1 gets an endowment which he or she can either share with player 2 or not. The second player does not have the ability to reject the offer made by the first player. In the *Ultimatum Game*, player 1 proposes how to divide the coins. Here an example of a fair offer and an unfair offer is depicted. In the fair and unfair offer the top row of coins are for Player 1, while the bottom row are coins for Player 2. Player 2 can either accept or reject the offer. When the offer is accepted both players get the number of coins as proposed, however, when the offer is rejected both players receive no coins. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174

Adolescents typically repay the trust a little less frequently than do adults, and children are the least trustworthy. Children may reciprocate in about 30% of cases and adolescents in 40% of cases. The percentage stabilizes at around 50% during early adulthood. This means that as adolescents get older, they more often take the perspective of the person dividing the money, and they become more prosocial—that is to say, more concerned about others. Possibly, they also get better at controlling the impulse to choose selfishly (van den Bos, Westenberg, Van Dijk, & Crone, 2010).

Dictator Game

In the Dictator Game, there also are two players who receive a certain sum of money (see Figure 5.3). The first player decides how to split this amount of money. The second player does not have the ability to reject the offer made by the first player. In other words, the proposed division is always divided as the first player suggests. The Dictator Game is thought to capture an objective indication of the fairness orientation of the first player because there are no strategies that play a role in the decision how to divide the sum of money (Van Dijk & Vermunt, 2000).

As it turns out, most people will give some money to the other person. They do not give necessarily half of their budget but usually 20% to 30% and keep the rest of the money for themselves. Children aged 8 to 10 years already do this. Children are inherently social from a young age; they also care about what other people get (Güroĝlu, van den Bos, & Crone, 2009).

Ultimatum Game

The Ultimatum Game is an economic exchange game played by two players, a proposer and a responder (see Figure 5.3; Güth, Schmittberger, & Schwarze, 1982). The proposer receives a given sum of money, the stake, and is asked to share the stake by offering a certain amount of the stake to the responder. If the responder accepts the offer, both players keep the amount allocated by the proposer. However, if the responder rejects the offer, both players go empty-handed. Based on economic rationality hypothesis (Von Neumann & Morgenstern, 1953), one would expect that responders accept all offers higher than nothing to maximize their personal gain. However, responders reject most offers lower than 30% of the share and exhibit a strong preference for fair offers hovering around a 50–50 split (Van Dijk, & Vermunt, 2000). This, too, is seen in young children—it seems that the sense of fairness is ingrained very early on. However, the intentions of the person making the offer make a difference. Young children take intentions into account less than adults do, and adolescents respond to intentions more than children do but less than adults do (Güroĝlu et al., 2009).

Neuroscience findings have offered exciting new perspectives on fairness considerations, and the developmental changes across childhood and adolescence. Again, we argue that cognitive control is a very important component in these processes.

Development of Fairness Perspectives

ERP studies have shown that receiving unfair offers compared to fair offers in an Ultimatum Game is associated with larger FRN. In these studies, the FRN was referred to as the MFN, however, for reasons of clarity, we continue to use the acronym "FRN" (Alexopoulos, Pfabigan, Lamm, Bauer, & Fishmeister, 2012; Boksem & De Cremer, 2010; Campanhã, Minati, Fregni, & Boggio, 2011; Polezzi et al., 2008; Van der Veen & Sahibdin, 2011; Wu, Zhou, van Dijk, Leliveld, & Zhou, 2011). The FRN not only differentiates between fair and unfair; the level of unfairness and individual differences also modulate the FRN. More specifically, the FRN is more pronounced for highly unfair offers than for moderately unfair offers (Hewig et al., 2011; Polezzi et al., 2008; Van der Veen & Sahibdin, 2011). Individual differences analysis has shown that, in particular, participants with a high concern for fairness show higher FRN responses to receiving unfair offers (Boksem & De Cremer, 2010).

Until recently, research has focused mainly on what happens when someone receives an unfair offer. However, in daily life, people also encounter situations in which they propose a fair deal but the other party does not accept this deal. In an ERP experiment, Zanolie, de Cremer, Güroğlu, and Crone (2015) tested whether the rejection of a fair offer would show a larger FRN compared to the acceptance of a fair offer or rejection of an unfair offer. Participants of two age groups, midadolescents 14 to 17 years old and early adults 19 to 24 years old, played an adapted version of the Ultimatum Game with another (computer-simulated) player. The participant always was the proposer and played multiple rounds with the same other player, by making a choice between an unfair distribution (7 coins for proposer and 3 for responder; 7/3) and one of two alternatives: a fair distribution (5/5) or a hyperfair distribution (3/7). Zanolie et al. (2015) found that the rejection of a fair offer (5/5; alternative was 7/3) was associated with a larger FRN compared to acceptance of a fair offer *and* rejection of an unfair offer (7/3; alternative was 3/7). The neural responses did not differ for the two age groups, suggesting that the FRN reacts as an alarm system to social prediction errors, which is already in place during midadolescence.

This developing preference for fairness over self-interest across childhood is thought to depend in part on the acquisition of perspective-taking abilities, enabling children to take another person's view (Takagishi, Kameshima, Schug, Koizumi, & Yamagishi, 2010). This progressing ability to take the perspective of others, which is defined here as the ability to understand thoughts and intentions of others and willingness to act on this understanding, subsequently can result in the development of strategic behavior. In order to examine strategic behavior, the Ultimatum Game and Dictator Game have been used to look at the behavior of the proposer.

An earlier neuroimaging study examined the development of fairness considerations from the proposer perspective (Steinbeis, Bernhardt, & Singer, 2012). In this study, children and adolescents between the ages 6 to 13 years played both the Dictator Game and the Ultimatum Game as proposers. Given that the Ultimatum Game requires participants to think about possible rejections by the receiver (this is not the case for the Dictator Game), the differences between offers in the games were seen as indexes of strategic behavior. The researchers reported that when entering adolescence (between ages 6 and 13 years), participants more often made strategic choices. Moreover, an increase in strategic offers was associated with more activity in the DLPFC. According to the researchers, this finding indicates that the DLPFC is important for the control of our impulse to be self-centered.

Interestingly, when a comparison is made between receiving an unfair offer and a fair offer in a classic Ultimatum Game, adolescents and adults showed stronger activity in the lateral PFC when receiving unfair offers. Possibly, the older participants were better able to inhibit initial impulses to reject, and they may have thought more about why someone would make an unfair offer (Steinmann et al., 2014).

Finally, an fMRI study that manipulated the intentions for proposing unfair offers found that the lateral PFC was differentially activated across adolescent development when the other player had no option but proposing an unfair split. The adults activated the lateral PFC and the temporoparietal junction if the dividing player had no other option; however, young adolescents did not yet activate these brain areas. Adolescents are taking the other person's perspective more and more during this game, which requires players to consider what is good for them and what is good for someone else. Also, adolescents make more and more use of the PFC, the area of the brain that is important in directing behavior.

Taken together, research that made use of social dilemma paradigms combined with EEG or fMRI recordings provided evidence that besides social brain regions, regions that are typically associated with cognitive control (such as the DMPFC and lateral PFC) also were activated. More interestingly, these regions in particular showed developmental changes when making social decisions, suggesting that cognitive control development may explain at least part of the changes we observe in making social decisions.

FUTURE DIRECTIONS

This chapter provided a review of neuroscience methods to understand changes in the development of cognitive control. Moreover, we have summarized how these changes are important for understanding cognitive development and also affective and social development.

Given that these results are based on several different paradigms and methods, future studies should examine the role of cognitive control in a more detailed way, for example, by specifically manipulating cognitive control demands (e.g., by having participants perform a working memory task while making social decisions) or by training cognitive control and test the effects on socioaffective decision making (Steinbeis et al., 2012).

Other directions for future studies will be focusing more on the role of individual differences. In the studies described in this chapter, we did not take these into account, but some studies have reported that, for example, socioeconomic status has an effect on the development of response inhibition (Spielberg et al., 2015).

Finally, an important direction for future studies will be to use longitudinal methods. This is commonly done in behavioral research, but only recently researchers have started to use these methods in neuroimaging designs. This type of analyses will prove especially useful for fitting growth trajectories (Ordaz, Foran, Velanova, & Luna, 2013) and for prediction analyses (Dumontheil & Klingberg, 2012).

CONCLUSION

The last two decades have resulted in tremendous progress in understanding the neural correlates of the development of cognitive control. Studies have found that, across late childhood and adolescence, individuals become better at keeping information in working memory, inhibiting ongoing behavior, and adjusting behavior according to changing environmental demands. Interestingly, these changes are accompanied by differential recruitment of regions in the PFC and parietal cortex. Many studies reported that activation in task-relevant brain regions in adults (brain regions that correlate with successful performance) become increasingly active when children develop into adults. These studies also show that children and adolescents often show activation in adjacent areas in the PFC and show a more diffuse pattern of activation than adults do. One way this can be explained is by interactive specialization. This theory suggests that the role of certain cortical brain regions, and the way they respond to stimuli in the environment, is the result of interaction and competition between these regions to acquire their roles. Some brain regions may have broad functionality early on in development and are partially activated in a wide range of functions. During development, activity-dependent interactions make cortical regions become more specialized (Johnson, 2011).

However, besides these cognitive changes, a key aspect of adolescence is social reorientation. Adolescents become increasingly more independent from their parents and orient themselves toward their peers. Adolescence is a period in which friendships become more important, and adolescents start forming new complex networks of peers. In order to navigate complex social environments, adolescents need to develop their social competencies. Furthermore, adolescents often weigh short-term and long-term outcomes differently from adults.

We have reviewed studies that have shown that when adults make these complex affective (e.g., weighing short- and long-term outcomes and social fairness considerations) decisions, they recruit regions that are often associated with cognitive control, such as the PFC. Interestingly, these regions increasingly contribute to socioaffective decision making when children and adolescents grow up. Obviously, many other brain regions play a role in affective decision making (e.g., the ventral striatum) and in social decision making (e.g., regions in the social brain network). Nonetheless, we have provided evidence that cognitive control is also very important in making socioaffective choices, and both EEG and neuroimaging methods provide valuable methods to understand these contributions in more detail.

REFERENCES

Achterberg, M., Peper, J. S., Van Duijvenvoorde, A. C., Mandl, R. C., & Crone, E. A. (2016). Fronto-striatal white matter integrity predicts development in delay of gratification: a longitudinal study. *Journal of Neuroscience*, *36*, 1954–1961.

Alexopoulos, J., Pfabigan, D. M., Lamm, C., Bauer, H., & Fischmeister, F. P. (2012). Do we care about the powerless third? An ERP study of the three-person ultimatum game. *Frontiers in Human Neuroscience*, *6*, 59.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Review Neuroscience*, *7*, 268–277.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in Cognitive Sciences*, *18*(4), 177–185.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 8, pp. 47–89). New York, NY: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). *Working Memory: The Multiple Component Model*. New York, NY: Cambridge University Press.

Banich, M. T., De La Vega, A., Andrews-Hanna, J. R., Mackiewicz Seghete, K., Du, Y., & Claus, E. D. (2013). Developmental trends and individual differences in brain systems involved in intertemporal choice during adolescence. *Psychology of Addictive Behaviors*, *27*(2), 416–430.

Barcelo, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia*, *40*(3), 349–356.

Barrouillet, P. (1996). Transitive inferences from set-inclusion relations and working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 1408–1422.

Bayliss, D. M., Jarrold, C., Baddeley, A. D., & Gunn, D. M. (2005). The relationship between short-term memory and working memory: complex span made simple? *Memory*, *13*, 414–421.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*, 122–142.

Blakemore, S. J. (2008). The social brain in adolescence. *Nature Review Neuroscience*, *9*, 267–277.

Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207.

Boksem, M. A., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, *5*(1), 118–128.

Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Trommer, B. L., Davenport, N. D., . . . Mesulam, M. M. (2003). Neural development of selective attention and response inhibition. *NeuroImage*, *20*(2), 737–751.

Brahmbhatt, S. B., White, D. A., & Barch, D. M. (2010). Developmental differences in sustained and transient activity underlying working memory. *Brain Research*, *1354*, 140–151.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, *33*(2), 301–311.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology*, *17*, 243–250.

Campanhã, C., Minati, L., Fregni, F., & Boggio, P. S. (2011). Responding to unfair offers made by a friend: neuroelectrical activity changes in the anterior medial prefrontal cortex. *The Journal of Neuroscience*, *31*, 15569–15574.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 367–379.

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*, 295–319.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, *54*, 241–257.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3), 104–110.

Cauffman, E., Shulman, E. P., Steinberg, L., Claus, E., Banich, M. T., Graham, S., & Woolard, J. (2010). Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. *Developmental Psychology*, *46*(1), 193–207.

Christakou, A., Brammer, M., & Rubia, K. (2011). Maturation of limbic corticostriatal activation and connectivity associated with developmental changes in temporal discounting. *NeuroImage*, *54*(2), 1344–1354.

Ciesielski, K. T., Lesnik, P. G., Savoy, R. L., Grant, E. P., & Ahlfors, S. P. (2006). Developmental neural networks in children performing a Categorical N-Back Task. *NeuroImage*, *33*(3), 980–990.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*, 636–650.

Crone, E. A., Ridderinkhof, K. R., Worm, M., Somsen, R. J. M., & Van der Molen (2004). Switching between spatial stimulus-response mappings: a developmental study of cognitive flexibility. *Developmental Science*, *7*, 443–455.

Crone, E. A., & van der Molen, M. W. (2004). Developmental changes in real life decision making: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology*, *25*(3), 251–279.

Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L., & Bunge, S. A. (2006). Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Acadamy of Sciences of the United States of America*, *103*(24), 9315–9320.

Crone, E. A., Zanolie, K., Van Leijenhorst, L., Westenberg, P. M., & Rombouts, S. A. R. B. (2008). Neural mechanisms supporting flexible performance adjustment during development. *Cognitive, Affective, & Behavioral Neuroscience*, *8*, 165–177.

Crowley, M. J., Wu, J., Crutcher, C., Bailey, C. A., Lejuez, C. W., & Mayes, L. C. (2009). Risk-taking and the feedback negativity response to loss among at-risk adolescents. *Developmental Neuroscience*, *31*, 137–148.

Crowley, M. J., Wu, J., Hommer, R. E., South, M., Molfese, P. J., Fearon, R. M. P., & Mayes, L. C. (2013). A developmental study of the feedback-related negativity from 10–17 years: age and sex effects for reward versus non-reward. *Developmental Neuropsychology*, *38*, 595–612.

Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *24*, 1742–1752.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142.

D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain and Cognition*, *41*, 66–86.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 466–503). New York, NY: Oxford University Press.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*, 331–338.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, *22*(5), 1078–1085.

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., & Casey, B. J. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science*, *9*(1), 1–8.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short term memory, and general fluid intelligence: a latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*, 1043–1053.

Fareri, D. S., & Delgado, M. R. (2014). Differential reward responses during competition against in- and out-of-network others. *Social Cognitive and Affective Neuroscience*, *9*, 412–420.

Flagan, T., & Beer, J. S. (2013). Three ways in which midline regions contribute to self-evaluation. *Frontiers in Human Neuroscience*, *7*, 450.

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: response to commentary. *Human Brain Mapping*, *32*, 2267–2269.

Geier, C. F., Garver, K., Terwilliger, R., & Luna, B. (2009). Development of working memory maintenance. *Journal of Neurophysiology*, *101*(1), 84–99.

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self-and externally caused failure of goal achievement. *NeuroImage*, *47*, 2023–2030.

Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences*, *1021*, 77–85.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., . . . Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 8174–8179.

Gunther Moor, B. G., van Leijenhorst, L., Rombouts, S. A., Crone, E. A., & Van der Molen,

M. W. (2010). Do you like me? Neural correlates of social evaluation and developmental trajectories. *Social Neuroscience*, *5*, 461–482

Güroĝlu, B., van den Bos, W., & Crone, E. A. (2009). Fairness considerations: increasing understanding of intentionality during adolescence. *Journal of Experimental Child Psychology*, *104*(4), 398–409.

Grammer, J. K., Carrasco, M., Gehring, W. J., & Morrison, F. J. (2014). Age-related changes in error processing in young children: a school based investigation. *Developmental Cognitive Neuroscience*, *9C*, 93–105.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*, 367.

Guyer, A. E., Choate, V. R., Pine, D. S., & Nelson, E. E. (2012). Neural circuitry underlying affective response to peer feedback in adolescence. *Social, Cognitive, & Affective Neuroscience*, *7*, 81–92.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*, 148–154.

Hämmerer, D., Li, S.-C., Müller, V., & Lindenberger, U. (2011). Life span differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *Journal of Cognitive Neuroscience*, *23*, 579–592.

Hester, R., Fassbender, C., & Garavan, H. (2004). Individual differences in error processing: a review and reanalusis of three event-related fMRI studies using the GO/NOGO task. *Cerebral Cortex*, *14*, 986–994.

Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M., Holroyd, C. B., & Miltner, W. H. (2011). Why humans deviate from rational choice. *Psychophysiology*, 48, 507–514.

Hitch, G. J., Towse, J. N. & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, *130*, 184–198.

Holroyd, C. B., Baker, T. E., Kerns, K. A., & Müller, U. (2008). Electrophysiological evidence of atypical motivation and reward processing in children with attention-deficit hyperactivity disorder. *Neuropsychologia*, *46*, 2234–2242.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*, 497–498.

Hooper, C. J., Luciana, M., Conklin, H. M., & Yarger, R. S. (2004). Adolescents' performance on the Iowa Gambling Task: implications for the development of decision making and ventromedial prefrontal cortex. *Developmental Psychology*, *40*(6), 1148–1158.

Huettell, S. A., Song, A. W., & McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*, *3rd Edition*. Sunderland, MA: Sinauer Associates.

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia*, *44*, 2017–2036

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21.

Jolles, D. D., Kleibeuker, S. W., Rombouts, S. A., & Crone, E. A. (2011). Developmental differences in prefrontal activation during working memory maintenance and manipulation for different memory loads. *Developmental Science*, *14*(4), 713–724.

Kerns, J. G., Cohen, J. D., MacDonald, A. W. III, Cho, R.Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*, 1023–1026.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317–324.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience*, *14*(1), 1–10.

Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Acadamy of Sciences of the United States of America*, *99*(20), 13336–13341.

Ladouceur, C. D., Dahl, R. E., & Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Developmental Science*, *10*, 874–891.

Lee, H. W., Lo, Y. H., Li, K. H., Sung, W. S., & Juan, C. H. (2015). The relationship between the development of response inhibition and intelligence in preschool children. *Frontiers in Psychology*, *6*, 802.

Libertus, M. E., Brannon, E. M., & Pelphrey, K. A. (2009). Developmental changes in category-specific brain responses to numbers and letters in a working memory task. *NeuroImage*, *44*(4), 1404–1414.

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and Cognition*, *72*, 101–113.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503–507.

Mills, K. L., Lalonde, F., Clasen, L. S., Giedd, J. N., & Blakemore, S. J. (2014). Developmental changes in the structure of the social brain in late childhood and adolescence. *Social, Cognitive, & Affective Neuroscience*, *9*(1), 123–131.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology*, *9*, 100–110.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

O'Doherty, J. P. (2011). Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Annuals of the New York Academy of Sciences*, *1239*, 118–129.

O'Hare, E. D., Lu, L. H., Houston, S. M., Bookheimer, S. Y., & Sowell, E. R. (2008). Neurodevelopmental changes in verbal working memory load-dependency: an fMRI investigation. *NeuroImage*, *42*(4), 1678–1685.

Olson, E. A., Collins, P. F., Hooper, C. J., Muetzel, R., Lim, K. O., & Luciana, M. (2009). White matter integrity predicts delay discounting behavior in 9- to 23-year-olds: a diffusion tensor imaging study. *Journal of Cognitive Neuroscience*, *21*(7), 1406–1421.

Ordaz, S. J., Foran, W., Velanova, K., & Luna, B. (2013). Longitudinal growth curves of brain function underlying inhibitory control through adolescence. *Journal of Neuroscience*, *33*(46), 18109–18124.

Owen, A. M., Evans, A. C., & Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. *Cerebral Cortex*, *6*, 31–38.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*, 46–59.

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience*, *9*, 947–957.

Peper, J. S., Mandl, R. C., Braams, B. R., de Water, E., Heijboer, A. C., Koolschijn, P. C., & Crone, E. A. (2013). Delay discounting and frontostriatal fiber tracts: a combined DTI and MTR study on impulsive choices in healthy young adults. *Cerebral Cortex*, *23*(7), 1695–1702.

Petanjek, Z., Judaš, M., Šimić, G., Rašin, M. R., Uylings, H. B. M., Rkic, P., & Kostović, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 13281–13286.

Peters, S., Braams, B. R., Raijmakers, M. E., Koolschijn, P. C., & Crone, E. A. (2014).

The neural coding of feedback learning across child and adolescent development. *Journal of Cognitive Neuroscience*, *26*, 1705–1720.

Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., & Rumiati, R. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research*, *190*(2), 218–223.

Richardson, C., Anderson, M., Reid, C. L., & Fox, A. M. (2011). Neural indicators of error processing and intraindividual variability in reaction time in 7- and 9-year-olds. *Developmental Psychobiology*, *53*, 256–265

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*, 443–447.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, *62*, 23–48.

Rubia, K., Lim, L., Ecker, C., Halari, R., Giampietro, V., Simmons, A., . . . Smith, A. (2013). Effects of age and gender on neural networks of motor response inhibition: from adolescence to mid-adulthood. *NeuroImage*, *83*, 690–703.

Rubia, K., Smith, A. B., Taylor, E., & Brammer, M. (2007). Linear age-correlated functional development of right inferior fronto-striato-cerebellar networks during response inhibition and anterior cingulate during error-related processes. *Human Brain Mapping*, *28*(11), 1163–1177.

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., & Brammer, M. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Human Brain Mapping*, *27*(12), 973–993.

Sakai, K., & Passingham, R. E. (2002). Prefrontal interactions reflect future task operations. *Nature Neuroscience*, *6*, 75–81.

Santesso, D. L., Dzyundzyak, A., & Segalowitz, S. J. (2011). Age, sex and individual differences in punishment sensitivity: factors influencing the feedback-related negativity. *Psychophysiology*, *48*, 1481–1489.

Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Developmental Science*, *9*, 473–481.

Satterthwaite, T. D., Wolf, D. H., Erus, G., Ruparel, K., Elliott, M. A., Gennatas, E. D., . . . Gur, R. E. (2013). Functional maturation of the executive system during adolescence. *Journal of Neuroscience*, *33*(41), 16249–16261.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*, 1835–1842.

Schel, M. A., & Crone, E. A. (2013). Development of response inhibition in the context of relevant versus irrelevant emotions. *Frontiers in Psychology*, *4*, 383.

Schel, M. A., Scheres, A., & Crone, E. A. (2014). New perspectives on self-control development: highlighting the role of intentional inhibition. *Neuropsychologia*, *65*, 236–246.

Scherf, K. S., Sweeney, J. A., & Luna, B. (2006). Brain basis of developmental change in visuospatial working memory. *Journal of Cognitive Neuroscience*, *18*(7), 1045–1058.

Segalowitz, S. J., Santesso, D. L., & Jetha, M. K. (2010). Electrophysiological changes during adolescence: a review. *Brain and Cognition*, *72*(1), 86–100.

Segalowitz, S. J., Santesso, D. L., Murphy, T. I., Homan, D., Chantziantoniou, D. K., & Khan, S. (2010). Retest reliability of medial frontal negativities during performance monitoring. *Psychophysiology*, *47*, 260–270.

Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: correspondence between hormonal and physical development. *Child Development*, *80*, 327–337.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661.

Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, *9*, 1007–1008.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcomes, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, *24*(38), 8223–8231.

Spencer-Smith, M., Ritter, B. C., Murner-Lavanchy, I., El-Koussy, M., Steinlin, M., & Everts, R. (2013). Age, sex, and performance influence the visuospatial working memory network in childhood. *Developmental Neuropsychology*, *38*(4), 236–255

Spielberg, J. M., Galarce, E. M., Ladouceur, C. D., McMakin, D. L., Olino, T. M., Forbes, E. E., . . . Dahl, R. E. (2015). Adolescent development of inhibition as a function of SES and gender: converging evidence from behavior and fMRI. *Human Brain Mapping*, *36*(8), 3194–3203

Steinbeis, N., Bernhardt, B. C., & Singer, T. (2012). Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron*, *73*(5), 1040–1051.

Steinbeis, N., Haushofer, J., Fehr, E., & Singer, T. (2016). Development of behavioral control and associated vmPFC-DLPFC connectivity explains children's increased resistance to temptation in intertemporal choice. *Cerebral Cortex*, *26*(1), 32–42.

Steinmann, E., Schmalor, A., Prehn-Kristensen, A., Wolff, S., Galka, A., Mohring, J., . . . Siniatchkin, M. (2014). Developmental changes of neuronal networks associated with strategic social decision-making. *Neuropsychologia*, *56*, 37–46.

Swanson, H. L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Memory and Cognition*, *32*, 648–661.

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of Experimental Child Psychology*, *105*, 130–137.

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(10), 1231–1238.

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: a review of developmental changes in the error-related negativity and brain maturation. *Developmental Cognitive Neuroscience*, *6*, 1–13.

Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A., & Klein, D. N. (2012). Electrocortical and behavioral measures of response monitoring in young children during a Go/No-Go task. *Developmental Psychobiology*, *54*, 139–150.

van den Bos, W., Güroĝlu, B., van den Bulk, B. G., Rombouts, S. A., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience*, *3*, 52.

van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2014). Connectivity strength of dissociable striatal tracts predict individual differences in temporal discounting. *Journal of Neuroscience*, *34*(31), 10298–10310.

van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Acadamy of Sciences of the United States of America*, *112*(29), E3765–E3774.

van den Bos, W., Westenberg, M., Van Dijk, E., & Crone, E. A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, *25*, 90–102.

van den Wildenberg, W. P., & van der Molen, M. W. (2004). Developmental trends in simple and selective inhibition of compatible and incompatible responses. *Journal of Experimental Child Psychology*, *87*(3), 201–220.

Van der Veen, F. M., & Sahibdin, P. P. (2011). Dissociation between medial frontal negativity and cardiac responses in the ultimatum game: effects of offer size and fairness. *Cognitive, Affective, & Behavioral Neuroscience*, *11*, 516–525.

Van Dijk, E., & Vermunt, R. (2000). Strategy and fairness in social decision making: sometimes it pays to be powerless. *Journal of Experimental Social Psychology*, *36*, 1–25.

van Duijvenvoorde, A. C. K., Zanolie, K., Rombouts, S. A. R. B., Raijmakers, M., & Crone, E.A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *Journal of Neuroscience*, *28*, 9495–9503.

van Meel, C. S., Heslenfeld, D. J., Rommelse, N. N., Oosterlaan, J., & Sergeant, J. A. (2012). Developmental trajectories of neural mechanisms supporting conflict and error processing in middle childhood. *Developmental Neuropsychology*, *37*, 358–378.

van Meel, C. S., Oosterlaan, J., Heslenfeld, D. J., & Sergeant, J. A. (2005). Telling good from bad news: ADHD differentially affects processing of positive and negative feedback during guessing. *Neuropsychologia*, *43*, 1946–1954.

van Noordt, S. J., & Segalowitz, S. J. (2012). Performance monitoring and the medial prefrontal cortex: a review of individual differences and context effects as a window on self-regulation. *Frontiers in Human Neuroscience*, *6*, 197.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, *30*, 829–858.

Vink, M., Zandbelt, B. B., Gladwin, T., Hillegers, M., Hoogendam, J. M., van den Wildenberg, W. P., . . . Kahn, R. S. (2014). Frontostriatal activity and connectivity increase during proactive inhibition across adolescence and early adulthood. *Human Brain Mapping*, *35*(9), 4415–4427

Von Neumann, J., & Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 255–274.

Wagner, A. D., Maril, A., Bjork, R. A., & Schacter, D. L. (2001). Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral prefrontal cortex. *NeuroImage*, *14*, 1337–1347.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*, 655–684.

Wu, Y., Zhou, Y., van Dijk, E., Leliveld, M. C., & Zhou, X. (2011). Social comparison affects brein responses to fairness in asset division: an ERP study with the ultimatum game. *Frontiers in Human Neuroscience*, *5*, 1–8.

Zanolie, K., De Cremer, D., Güroğlu, B., & Crone, E. A. (2015). Rejection in bargaining situations: An event-related potential study in adolescents and adults. *PLoS One*, *10*, e0139953.

Zanolie, K., Teng, S., Donohue, S. E., Band, G. P. H., Rombouts, S. A. R. B., & Crone, E. A. (2008). Switching between colors and shapes on the basis of positive and negative feedback: an fMRI and EEG study on feedback-based learning. *Cortex*, *44*, 537–547.

Zanolie, K., Van Leijenhorst, L., Rombouts, S., & Crone, E. (2008). Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia*, *46*(1), 117–126.

Zottoli, T. M., & Grose-Fifer, J. (2012). The feedback-related negativity (FRN) in adolescents. *Psychophysiology*, *49*, 413–420.

CHAPTER 6

Development of Mathematical Reasoning

TERESA IUCULANO AND VINOD MENON

INTRODUCTION

The term "mathematics" derives from the Greek word *máthēma* (i.e., "to learn"). The etymology of the word captures a key aspect of this discipline: Mathematical knowledge is hierarchical in nature, and learning of skills in a new domain builds on a lower-level primitive. The hierarchical structure underlying the most complex mathematical reasoning abilities is built on a fundamental system for representing numerosity ("number sense") (Dantzig, 1930; Dehaene, 1997): All forms of mathematical reasoning require fundamental knowledge of the basic properties of numbers, principles of cardinality and numerosity as abstract representations of sets, and the axiomatic rules by which quantity is manipulated.

Knowledge of numerical magnitude and manipulations of symbolic and nonsymbolic quantity (number sense) are critical building blocks from which all mathematical knowledge is "constructed" in the brain. These basic building blocks rely on visual and auditory association cortices, which help decode the visual form and phonological features of numerical stimuli, and the parietal attention system (Dehaene, Piazza, Pinel, & Cohen, 2003), which helps to build semantic representations of quantity (Ansari, 2008) from multiple low-level visuospatial primitives, notably the ability to attend to and individuate individual objects in space. These primitives are anchored in the posterior parietal cortex (PPC), particularly in its intraparietal sulcus (IPS) subdivision, and are engaged early in infancy many years before a child learns to process culturally determined numerical symbols and number words. As a child begins to learn the use of orthographic or phonological symbols, such as Arabic-Hindu numerals or number words, new representations develop in the fusiform gyrus (FG) in the ventral temporal occipital cortex (VTOC) and are mapped onto appropriate quantity representations (Allison, McCarthy, Nobre, Puce, & Belger, 1994; Ansari, 2008; Dehaene et al., 2004; Park, Hebrank, Polk, & Park, 2012; Shum et al., 2013). Next, procedural and working memory systems anchored in the basal ganglia and fronto-parietal circuits help create short-term representations that support the manipulation of multiple discrete quantities over several seconds. Finally, episodic and semantic memory systems play an important role in long-term memory formation and generalization beyond individual problem attributes; and prefrontal control processes guide and maintain attention in the service of goal-directed decision making. Hence, mathematical skills rely on distinct yet interacting neurocognitive processing systems. (See Figure 6.1.)

The way in which these neurocognitive systems are engaged depends critically on

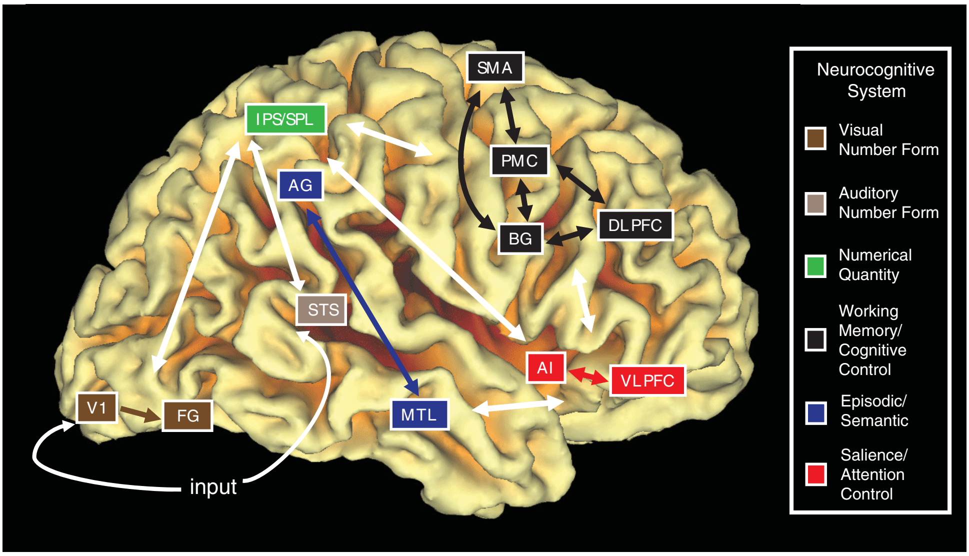


Figure 6.1 Schematic circuit diagram of brain regions involved in mathematical reasoning. The fusiform gyrus (FG) in the ventral occipital cortex (VTOC) decodes visual number form and together with the intraparietal sulcus (IPS) in the posterior parietal cortex (PPC)—which helps build visuospatial representations of numerical quantity (shown in green boxes and links)—form the building blocks of mathematical reasoning. The superior temporal gyrus (STG) aids in decoding auditory number words. Multiple parietal-frontal circuits link the IPS with working memory and cognitive control systems that include the dorsolateral prefrontal cortex (DLPFC), premotor cortex (PMC), supplementary motor area (SMA), and basal ganglia (BG). These circuits facilitate visuospatial working memory for objects in space and create a hierarchy of short-term representations that allow manipulation of multiple discrete quantities over short periods of time (i.e., several seconds). The declarative memory system anchored in the medial temporal cortex (MTL)—and particularly the hippocampus—plays an important role in long-term memory formation and generalization beyond individual problem attributes. Finally, prefrontal control circuits (shown in red) anchored in the anterior insula (AI) and ventrolateral prefrontal cortex (VLPFC) serve as flexible hubs for integrating information across attentional and memory systems, thereby facilitating goal-directed problem solving and decision making during mathematical reasoning. Relative transparency for BG and MTL indicates subsurface cortical structure. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
Source: Adapted from Fias, Menon, & Szucs, 2013, © 2013 by Elsevier.

both problem complexity and stage of an individual's cognitive development. This aspect is evident mostly in developmental studies that show that children's gains in problem-solving skills are characterized by shifts in the mix of problem-solving strategies, with inefficient procedural strategies gradually being replaced with direct retrieval of domain-relevant facts (Geary, Bow-Thomas, & Yao, 1992; Geary, Hoard, Byrd-Craven, & Desoto, 2004). Over time and with development, episodic and semantic memory systems build representations in long-term memory that allow for fast access of learned arithmetic facts (Butterworth, 1999; Geary, 1994; Menon, 2014; Siegler, 1996).

This chapter synthesizes emerging findings on the perceptual and cognitive building blocks of numerical cognition, the functional brain circuits associated with them, and the multiple memory and cognitive control systems that play a critical role in scaffolding children's mathematics learning and skill development. Consistent with the view of hierarchical learning, we provide evidence that the IPS and FG continue to play an important role in the development of more complex mathematical reasoning skills and that, with increased proficiency, they continue to anchor a "core" visuospatial number system for mathematical reasoning. We emphasize that these regions do not, however, function in isolation. They receive inputs from multiple brain regions from both visual and auditory association cortices involved in symbolic and number word processing and send outputs to several others. The development of core brain systems for mathematical reasoning is thus supported by multiple distributed neural processes involved in number form, magnitude and quantity representations, working memory, as well as cognitive control and declarative memory (see Figure 6.1) (Arsalidou & Taylor, 2011; Fias, Menon, & Szucs, 2013; Qin et al., 2014). We take a developmental systems neuroscience approach to shed light onto multiple functional circuits that mature as a child learns to reason "mathematically."

BUILDING BLOCKS OF MATHEMATICAL COGNITION

Nonsymbolic "Number Sense"

There is general consensus that humans, and other species, are endowed with a core capacity to represent numerosity and that this capacity may be innate (Butterworth, 1999; Carey, 2004; Dehaene, 1997). Infants can discriminate displays with different numerosity—they respond when the display changes from 2 objects to 3 or from 3 objects to 2 (Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1990; Van Loosbroek & Smitsman, 1990). Furthermore, infants respond to changes in the numerosity of a set, even when these changes take place behind a screen, suggesting a sensitivity to numerical transformations that underlies basic arithmetic processes (Wynn, 1992, 1998). In addition, visual perception of quantity is susceptible to adaptation (Burr & Ross, 2008), which suggests that numerosity can be considered as a primary visual property of a scene, much like color. An endowed capacity for numerosity may have evolutionary origins as numerical discrimination abilities have been found in a number of species, including orangutans (Shumaker, Palkovich, Beck, Guagnano, & Morowitz, 2001), monkeys (Brannon & Terrace, 1998), rats (Church, 1984; Mechner, 1958; Meck & Church, 1983), birds (Emmerton, Lohmann, & Niemann, 1997; Koehler, 1951), and even bees (Dacke & Srinivasan, 2008) and fish (Agrillo, Dadda, & Bisazza, 2007; Agrillo, Dadda, Serena, & Bisazza, 2008; Piffer, Agrillo, & Hyde, 2012). This finding points to the existence of a phylogenetically conserved ability to represent and process numerical magnitude. However, in humans, this core ability has undergone tremendous evolutionary expansion as a result of the availability of multiple functional brain circuits that children can draw on to guide the development of complex numerical cognitive abilities.

Theories of the development of numerical cognition in humans posit two subsystems of core knowledge: One dealing with small quantities, referred to as an object file or subitizing system (Cantlon, Platt, & Brannon, 2009b; Carey, 2004; Feigenson, Dehaene, & Spelke, 2004; Hyde, Boas, Blair, & Carey, 2010; Le Corre & Carey, 2007) (see Figure 6.2A), while the other represents larger quantities and is referred to as the approximate number system (ANS) (Dehaene, 2003; Gebuis, Cohen Kadosh, & Gevers, 2016) (see Figure 6.2B). This latter system comes into play when there are

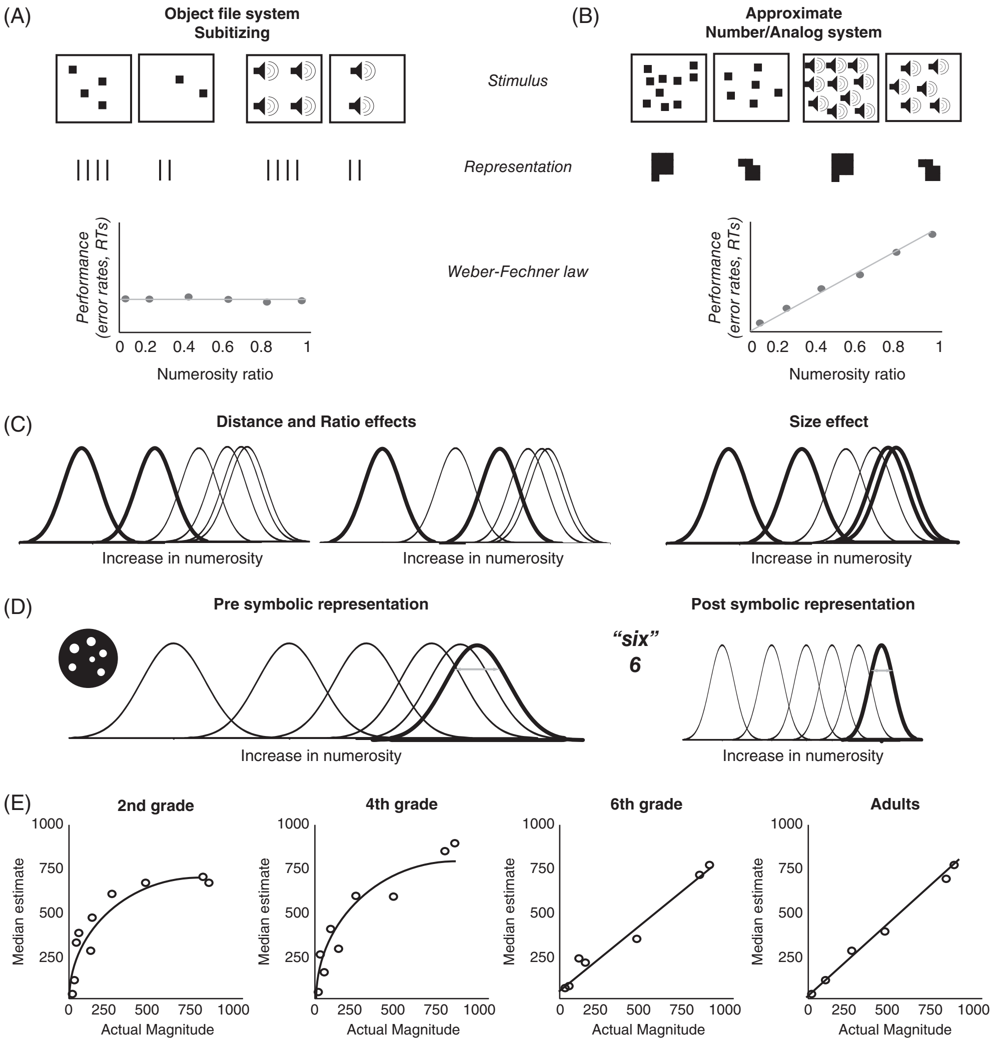


Figure 6.2 Building blocks of mathematical reasoning. (A) The object file (or subitizing) system deals with up to 4 items in the subitizing range, and it is not subject to the Weber-Fechner law, so that all items are represented similarly. (B) The approximate number/analog system deals with quantities greater than 4 and creates analog representations of numerosity that are subject to the Weber-Fechner law similar to any physical stimulus that varies along a continuum. (C) Schematic representation of numerosity on a compressed logarithmic number line. *Distance and ratio effects*: Numerosity representations show less prominent overlap with increasing distance (and ratio) between them. As a result, the discriminability increases with numerical distance. *Size effect*: Numerosity representations show greater overlap with set size. As a result, it becomes harder to compare larger number pairs, even when the distance between them is kept constant. (D) The width of the Gaussian curves representing quantity becomes smaller with increased proficiency, reflecting maturation of abstract symbolic representations. (E) Fine-tuning of number representations from early years of schooling to adulthood. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
Source: Panel E is adapted from Siegler & Opfer, 2003, © 2003 by SAGE.

more than 5 items to be tracked by the object-file/subitizing system (Feigenson et al., 2002). The object-file/subitizing system creates exact representations of a small set of items (i.e., fewer than 4) and is highly limited in capacity (Figure 6.2A). In contrast, the ANS can flexibly represent numbers much greater than 5, albeit with increasingly poor resolution along a logarithmically compressed number line (Figure 6.2B–C). The ANS has several key properties, including sensitivity to size and distance between quantities. It is harder to discriminate two sets of items as set sizes increase in magnitude. For example, it is more difficult to compare 8 versus 9 than it is to compare 3 versus 4 (*size effect*) (Figure 6.2C, right panel). Performance, assessed by accuracy and reaction time, drops dramatically when the quantities to be compared are closer in magnitude, so that it is more difficult to compare 3 versus 4 than it is to compare 3 versus 9 (*distance* and *ratio effects*) (Figure 6.2C, left panel). Consistent with the Weber-Fechner law (Fechner, 1860), the more quantity representations overlap, the more difficult it is to discriminate them (Figure 6.2C–D left panel). The acuity of the ANS is therefore best described by the relative difference (ratio) between the quantities to be compared (Figure 6.2A–B). The Weber-Fechner law holds true for any given stimuli that vary along physical continua, such as luminance, time, pitch tone, length, and others (Moyer & Landauer, 1967). Within this account, it has been proposed that numerosities are encoded and represented holistically as analog magnitudes (Figure 6.2B) (Feigenson, Carey, & Spelke, 2002; Feigenson et al., 2004; Leibovich, Katzin, Harel, & Henik, 2016).

The discrete and analog dimensions are correlated with one other and co-mature over development, making it difficult to disentangle their relative causal roles. Infants as young as 2 months do not respond to numerosity per se but rather to changes in continuous analog dimensions, such as surface area and contour length (Clearfield & Mix, 1999; Feigenson et al., 2002). Even children as old as 3 years have been shown to rely on nonnumerical visual cues in order to discriminate between sets of objects (Rousselle, Palmers, & Noel, 2004). With the acquisition and mastery of cultural symbols, such as Arabic-Hindu digits and number words, this system gains precision—the width of the Gaussian curves becomes smaller and "6-ness," for example, can be represented more exactly (Figure 6.2D, right panel). Overall, numerical representations become more linear and better tuned (Figure 6.2E) (Halberda & Feigenson, 2008; Siegler & Opfer, 2003), setting the stage for increased precision in mapping nonsymbolic quantity sets to their symbolic representations.

Symbolic Representations and Their Mapping to Nonsymbolic Representations

Mathematical skills are built upon understanding symbols and number words that represent quantity and facilitate their manipulation in ways that would be inefficient to do with nonsymbolic quantity. Numerical symbols facilitate the formation of precise numerosity representations and engender precise computational abilities that allow children to develop numerical skills that go far beyond those of any other primate. The symbolic systems for numbers in both the visual and auditory modalities share the same distinguishing properties of the two nonsymbolic systems for quantity knowledge outlined in the last section. There is evidence for a subitizing system that deals with small quantities in symbolic form, which might be due to the fact that the first number words are acquired initially for small cardinalities, 1 to 4 (Briars & Siegler, 1984; Carey, 2009), and only after the discovery of the counting principles are these number words extended to larger numerosities (Wynn, 1992, 1998). Thus, children's accuracy during a nonsymbolic estimation task depends critically on their counting abilities (Barth, Starr, & Sullivan, 2009).

Although the object file/subitizing system (Figure 6.2A) is unencumbered by the distance and ratio effects (Figure 6.2C) in children as well as adults, larger sets (Figure 6.2B) show distance/ratio and size effects similar to nonsymbolic sets (Figure 6.2C) (Piazza et al., 2010). In their seminal study, Moyer and Landauer (1967) demonstrated that when adults are asked to determine which of two numbers—presented visually as Arabic-Hindu digits or number words—is larger in magnitude, subjects were faster and less error prone when the ratio was high. Moreover, it has been demonstrated that symbolic distance/ratio effects are prominent in the context of priming, where the speed of processing of a target number is influenced directly by the numerical distance to a numerical stimulus presented immediately prior to the target (Reynvoet, Brysbaert, & Fias, 2002; Van Opstal, Gevers, De Moor, & Verguts, 2008). For example, when asked to judge if the target number "9" is smaller or larger than a standard number, such as "5," reaction times are faster if the preceding number is "2," which is 3 units from 5, compared to "4," which is 1 unit from 5. Critically, these effects are already present in children as young as 6 years (Reynvoet, De Smedt, & Van den Bussche, 2009). Together, this evidence suggests that an internal scale for symbolic number processing inherits some key properties of the nonsymbolic numerosity scale that governs preexisting and culturally independent representations of approximate quantities. Indeed, most existing proposals for the acquisition of symbolic numbers claim that the symbols for numbers acquire meaning by being mapped onto the preexisting, evolutionarily conserved, core nonsymbolic quantity representation system.

DEVELOPMENT OF ARITHMETIC PROBLEM-SOLVING SKILLS

Role of Foundational Symbolic and Nonsymbolic Skills

The ability to manipulate numerosity forms the basis of mathematical reasoning (Barth, La Mont, Lipton, & Spelke, 2005; Barth et al., 2006; Gilmore, McCarthy, & Spelke, 2007, 2010). Crucially, arithmetic problem-solving skills build on a core number knowledge system for representing numerical quantity using abstract symbols that is typically in place by the age of 5 (Barth et al., 2009; Lipton & Spelke, 2005). In order to perform exact manipulations, such as addition and subtraction, a child needs to develop fine-tuned representations of numbers (Butterworth, 2010). Thus, a child needs to be able to enumerate distinct items in a set and assess whether a set of items represents a greater or smaller quantity compared to another set: A set of 3 items is less than a set of 4 items. Next, it is essential to be able to manipulate numbers: A set of 3 items together with a set of 4 items forms a larger set of 7 items (Giaquinto, 1995). Furthermore, the ability to represent large quantities—and to manipulate them efficiently—depends on the acquisition of culture-specific numerical symbols, such as Arabic-Hindu digits and number words (Ansari, 2008). Indeed, higher math skills have been associated with better mapping abilities between the symbolic and nonsymbolic systems for numbers (Iuculano, Tang, Hall, & Butterworth, 2008; Lyons & Ansari, 2015; Rousselle & Noel, 2007).

Both symbolic and nonsymbolic core skills have been linked to mathematical abilities. Several studies have shown a significant relationship between the nonsymbolic ANS (Figure 6.2B) and mathematical skills (Inglis, Attridge, Batchelor, & Gilmore, 2011; Mazzocco, Feigenson, & Halberda, 2011; Starr, Libertus, & Brannon, 2013). In one study, performance on a standardized math test in 14-year-olds was retroactively predicted by their ability to accurately and precisely discriminate between nonsymbolic numerosity at a very early age, as early as kindergarten (Halberda, Mazzocco, & Feigenson, 2008). However, most studies have found this relation between nonsymbolic number sense and arithmetic skills only in children (i.e., 7- to 9-year-olds) but not in adults (Inglis et al., 2011). Others have not found such relationship, even in children, and have instead reported stronger relations with symbolic number sense (De Smedt, Verschaffel, & Ghesquiere, 2009; Gilmore et al., 2013; Holloway & Ansari, 2009; Iuculano et al., 2008; Rousselle & Noel, 2007; Soltesz, Szucs, & Szucs, 2010).

Verbal processes involved in learning the labels of Arabic-Hindu numerals and the ability to translate between Arabic-Hindu numerals and verbal codes also play a critical role in the development of mathematical reasoning (De Smedt, Taylor, Archibald, & Ansari, 2010; Gobel, Watson, Lervag, & Hulme, 2014). Like literacy, symbolic numeracy exists only in cultures where it is explicitly instructed. In order to develop concepts of exact numerosity, children initially depend on language-based number words (Carey, 2004) and their relation to visual symbols (Ansari, 2008). Hence, the development of verbal counting principles provides a critical foundation for later mathematical achievement (Butterworth, 2005; Gelman, 1978). Indeed, mastery of the how-to-count principles has been found to predict children's later mathematical reasoning abilities (De Smedt et al., 2010; Passolunghi, Vercelloni, & Schadee, 2007). Crucially, proficiency in mathematics is strongly associated with better mapping abilities between the symbolic and nonsymbolic systems for numbers (Barth et al., 2009; Iuculano et al., 2008; Mundy & Gilmore, 2009; Rousselle & Noel, 2007).

Working Memory and Cognitive Control

Working memory and associated cognitive control processes play an essential role in the development of children's problem-solving abilities (Geary & Widaman, 1992; Geary et al., 2004; Siegler, 1987, 1996). Indeed, the particular emphasis on working memory in mathematical cognition is most prominent in developmental studies. This observation has origins in children's immature problem-solving abilities, which require them to break down numerical problems into more basic components (Cowan et al., 2011; Wu et al., 2008). For example, compared to adults, children rely more on counting strategies during simple arithmetic problem solving (Qin et al., 2014), and they access multiple working memory components, including short-term storage, rule-based manipulation and updating of the stored contents (Ashcraft, 1992; Qin et al., 2014). With increased proficiency and a shift to fact retrieval strategies, there is less demand and need for working memory resources (Bailey, Littlefield, & Geary, 2012; Geary, 1994; Geary et al., 2007; Menon, 2014; Siegler, 1996).

Working memory processes used in mathematical reasoning draw on the object-file/subitizing system (Figure 6.2A), which plays an early developmental role in tracking items over a short period of time and in space, allowing manipulation of multiple discrete items over several seconds (Carey, 2004). More complex manipulations of quantities require the aid of higher-order cognitive functions that support maintenance of numerical representations in working memory over longer time periods spanning several seconds. Furthermore, in order to solve complex problems successfully, children need to learn to inhibit irrelevant information and implement task-relevant rules (e.g., "+" or "–" signs that characterize arithmetic operations). Crucially, recent longitudinal studies have highlighted the importance of the executive as well as the visual components of working memory in predicting math proficiency (Geary, 2011). Weak interference suppression has also been proposed to lie at the core of difficulties in acquiring age-appropriate mathematical abilities (Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013), either because of the inability to suppress intrusion errors during calculation (Geary & Menon, in press) or because of interference with the contents of working memory (Oberauer, Lange, & Engle, 2004; Unsworth & Engle, 2007). These domain-general cognitive processes are as vital as core numerical knowledge; they not only provide necessary scaffolds for the development of more efficient strategies during the initial stages of arithmetic learning and skill development (Bull, Espy, & Wiebe, 2008), but they also facilitate learning of new and more complex materials in adolescence and adulthood (Zamarian & Delazer, 2014).

The manner in which distinct components of working memory contribute to the development of mathematical problem-solving abilities is considered further in the section titled "Neurodevelopmental Processes in Mathematical Reasoning."

Development of Memory-Based Strategies

The ability to retrieve basic facts efficiently from memory is a vital aspect of mature problem solving (Siegler, 1996). During development, children's math problem-solving skills gradually become more dependent on memory-based strategies, such as direct retrieval, and less dependent on effortful procedures, such as counting (Geary et al., 2004; Qin et al., 2014). During learning and skill acquisition, children have available to them a mix of procedural and memory-based approaches for solving problems in a given domain (Siegler, 1996), whereby elaborate computational procedures (e.g., counting to solve arithmetic problems, phonetic decoding to sound out and read words) are used more frequently than memory-based procedures (e.g., addition fact or word retrieval) (Butterworth, 1999; Geary, 1994; Menon, 2014; Siegler, 1996). Less efficient strategies then are gradually replaced by more efficient memory-based ones (Geary, 1994; Siegler, 1996).

By age 7, children are typically able to answer single-digit addition problems, although rapid fact retrieval has not fully developed by this age (Jordan, Hanich, & Kaplan, 2003). Between ages 7 and 9, problem-solving abilities generally progress from effortful counting strategies to more automatic retrieval strategies, although the extent of such skill maturation and the sources of individual variability are less well understood. During successful mathematical learning, the representation of arithmetic facts in long-term memory is aided by the repeated use of counting and other updating procedures (Ashcraft, 1982; Siegler & Shipley, 1995; Siegler & Shrager, 1984). For example, when counting on from 3 to 7 to solve the problem "3+4," an association is dynamically formed between the correct solution ("7") and the addends of the problem ("3 and 4"). After many repetitions, children begin to directly retrieve the answer when presented with the problem (Siegler & Shrager, 1984).

In this manner, knowledge of counting principles, language, and working memory together support encoding of information into long-term semantic memory. In the section titled "Neurodevelopmental Processes in Mathematical Reasoning," we consider how interacting neurocognitive systems support these processes.

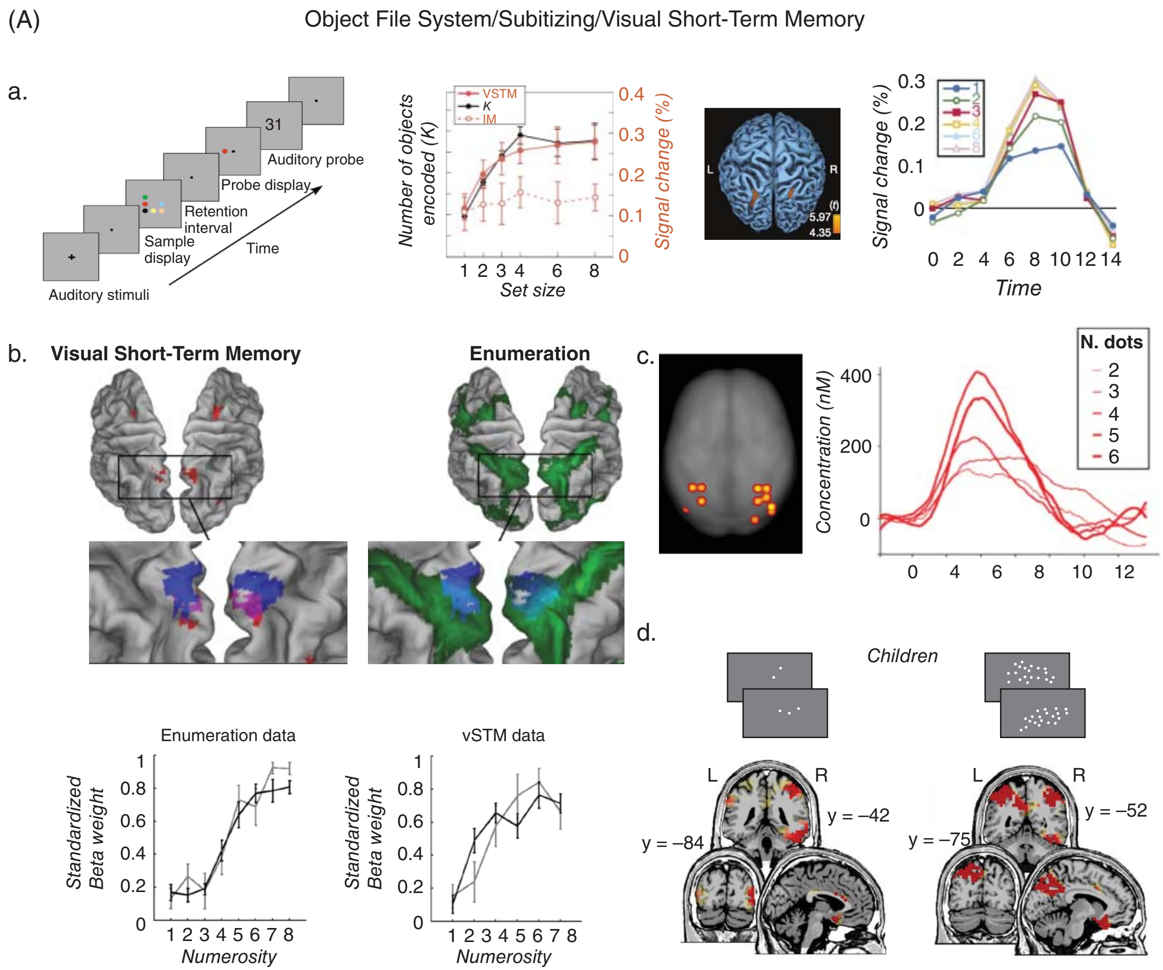
NEURODEVELOPMENTAL PROCESSES UNDERLYING BUILDING BLOCKS OF NUMERICAL COGNITION

Neural Building Blocks

The neural building blocks for numerical cognition are constructed from core hubs anchored in the IPS and the FG and their associated functional circuits (Figure 6.1). These hubs code perceptual and semantic representations of nonsymbolic and symbolic quantity and facilitate their dynamic manipulation in a context-dependent manner (Ansari, 2008). Much of the early evidence for the importance of the parietal lobe, including the IPS, in numerical cognition came from lesion studies in patients with brain damage. These studies date back to the seminal work of Henschen (1920) and Gerstmann (1940), which revealed that damage to the left inferior parietal cortex caused severe deficits in the ability to perform simple arithmetic calculations. In the ensuing decades, more refined neuropsychological investigations have determined that impairments in more basic functions, such as number comprehension and production, could also be traced to lesions of the inferior parietal cortex (Cipolotti, 2001; Dehaene & Cohen, 1997; Delazer & Benke, 1997; Lemer, Dehaene, Spelke, & Cohen, 2003). Beyond the inferior parietal cortex, difficulties in decoding numerical symbols have also been reported in a few individuals with ventral temporal occipital lesions, including the FG (Cohen & Dehaene, 1995).

Yet only with the advent of more precise functional brain mapping and neurodevelopmental studies have the functional roles of specific subdivisions of the PPC and the VTOC been clarified.

It is now established that functional systems associated with the IPS play a prominent role in mathematical cognition and its development. The reason for this is not that the IPS houses a number module, as previously suggested (Butterworth, 1999, 2010; Dehaene, 2003; Dehaene & Cohen, 1995), but that its connectivity with other brain areas helps sustain a broad class of processes, including spatial attention, individuation, and pointing of objects in extrapersonal space (Schaffelhofer & Scherberger, 2016; Simon, Mangin, Cohen, Le Bihan, & Dehaene, 2002) as well as the encoding of the locations of objects into visual short-term memory storage (see Figure 6.3A, a–b) (Knops, Piazza, Sengupta, Eger, & Melcher, 2014; Luck & Vogel, 2013; Todd & Marois, 2004). These functions play an important role in the individuation of objects in space, thereby contributing to the development of functional systems for the estimation of quantity (Cantlon & Brannon, 2006; Cantlon, Brannon, Carter, & Pelphrey, 2006; Cutini, Scatturin, Basso Moro, & Zorzi, 2014; Harvey, Klein, Petridou, & Dumoulin, 2013; Heine, Tamm, Wissmann, & Jacobs, 2011; Knops et al., 2014; Kucian, Von Aster, Loenneker, Dietrich, & Martin, 2008; Piazza et al., 2004; Vogel, Grabner, Schneider, Siegler, & Ansari, 2013) (Figures 6.3, 6.4A–B, 6.5A). Thus, for example, IPS activity is enhanced during visual short-term memory tasks in which participants are asked to remember the spatial location of objects (Todd & Marois, 2004) (Figure 6.3A, a). Critically, IPS activity increases with the number of target locations to be remembered, plateauing at about 4 items in close


(A)
Object File System/Subitizing/Visual Short-Term Memory
a.
Auditory stimuli
Sample display
Retention interval
Probe display
Auditory probe
31
Time
Number of objects encoded (K)
Signal change (%)
VSTM
K
IM
Set size
L
R
5.97
4.35
Signal change (%)
Time
b.
Visual Short-Term Memory
Enumeration
Enumeration data
vSTM data
Standardized Beta weight
Numerosity
c.
Concentration (nM)
N. dots
d.
Children
L
R
y = −84
y = −42
y = −75
y = −52

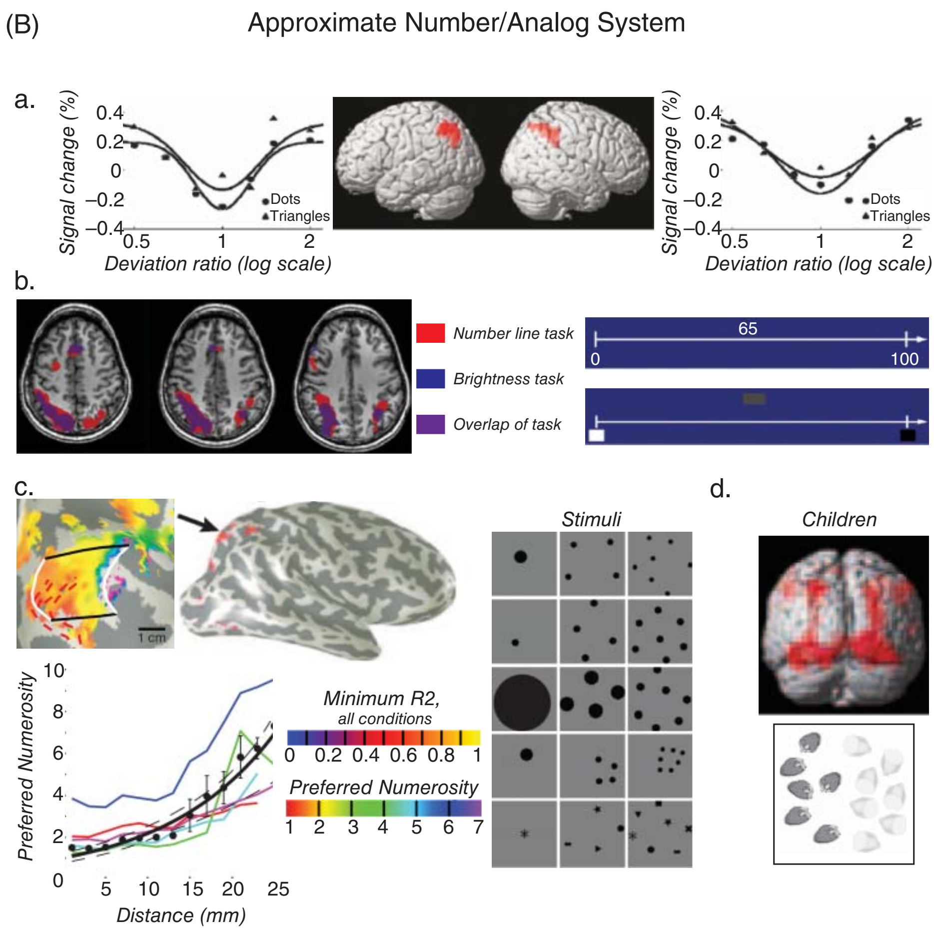


Figure 6.3 Neural building blocks of mathematical reasoning. (A) The visual short-term memory system (vSTM) in adults and children. (a) (left) vSTM task. Subjects judge whether the color of the probe matches the color of the disc shown at the same position in the same display; (center) behavioral performance and IPS response function shown as percentage of signal change in the vSTM task. Behavioral performance corresponding to the estimated number (K) of encoded color discs at each set size. (right) Time course of brain activation showing capacity limits in the IPS vSTM system. (b) Brain regions exhibiting a vSTM profile (red), and brain regions exhibiting an enumeration profile (green) with associated response profiles. (c) Hemodynamic response profiles in the right IPS related to enumeration of nonsymbolic quantity. (d) Brain activation in the parietal and occipital cortices for small (> 4) and large (< 4) numerosity in elementary school children as revealed by electroencephalogram recordings.
(B) The approximate number system in adults and children. (a) Ratio/distance effects in the bilateral parietal IPS in adults. (b) Brain regions involved in estimating the spatial position of numerical and other magnitudes overlap in the bilateral IPS in adults. (c) Numerosity-dependent functional gradients in the IPS. (d) Parietal activation during a nonsymbolic comparison task with larger numerosity in 9-year-old children.

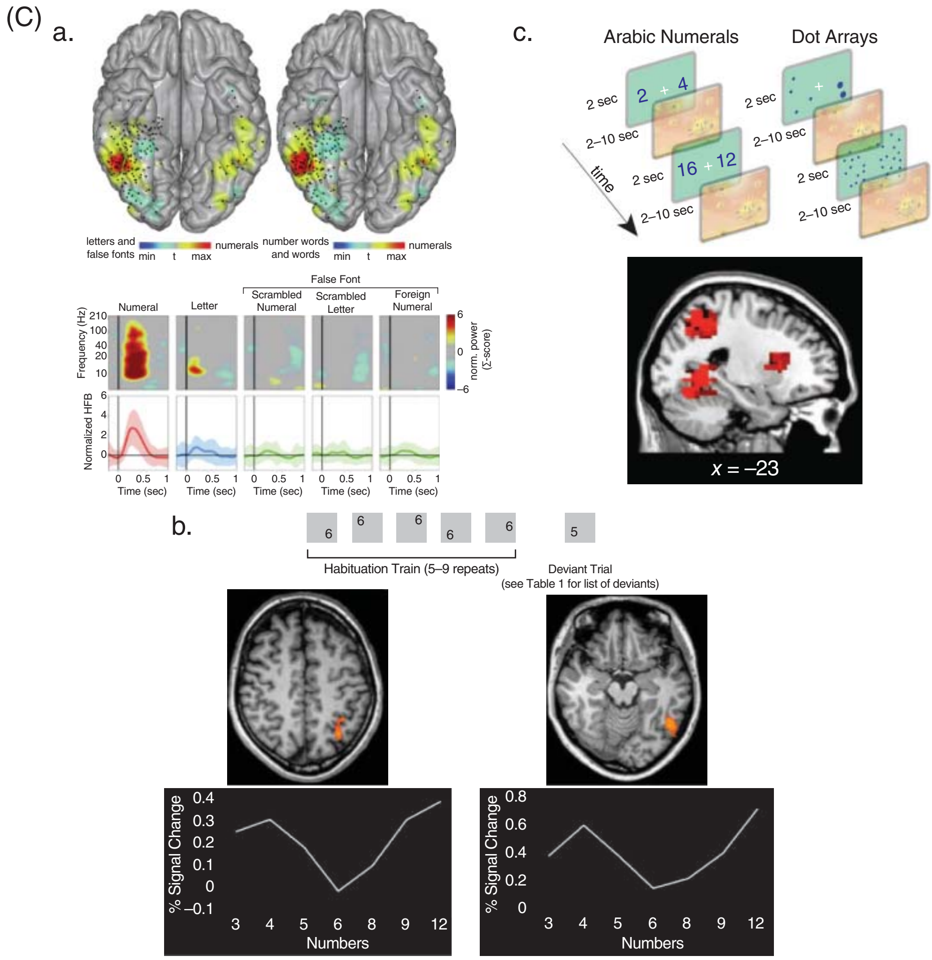


Figure 6.3 (C) (a) Visual number form in the fusiform gyrus (FG). Preferential gamma-band response to numerals in ventral temporal occipital cortex (VTOC). (b) Symbolic distance effects in the IPS and FG in adults. Left IPS and left FG show ratio-dependent modulation in response to Arabic digits in adults. (c) Sensitivity to symbolic and nonsymbolic quantities in 6-year-old children. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

Source: (A) (a) Adapted from Todd and Marois, 2004, © 2004 by NPG. (b) Adapted from Knops et al., 2014, © 2014 by SfN. (c) Adapted from Cutini et al., 2014, © 2014 by Elsevier. (d) Adapted from Heine et al., 2011, © 2011 by Elsevier.

(B) (a) Adapted from Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004, © 2004 by Elsevier. (b) Adapted from Vogel et al., 2013, © 2013 by Elsevier. (c) Adapted from Harvey et al., 2013, © 2013 by Science. (d) Adapted from Kucian et al., 2008, © 2008 by Taylor & Francis.

(C) (a) Adapted from Shum et al., 2013, © 2013 by SfN. (b) Adapted from Holloway, Battista, Vogel, & Ansari, 2013, © 2012 by the Massachusetts Institute of Technology. (c) Adapted from Cantlon et al., 2009a, © 2009 by the Massachusetts Institute of Technology.

correspondence with the subitizing range (Figure 6.2A). Regions of the posterior IPS, which are engaged during visual short-term memory tasks, overlap with those activated during enumeration tasks (Figure 6.3A, b) (Knops et al., 2014), further highlighting the close correspondence between the two functions.

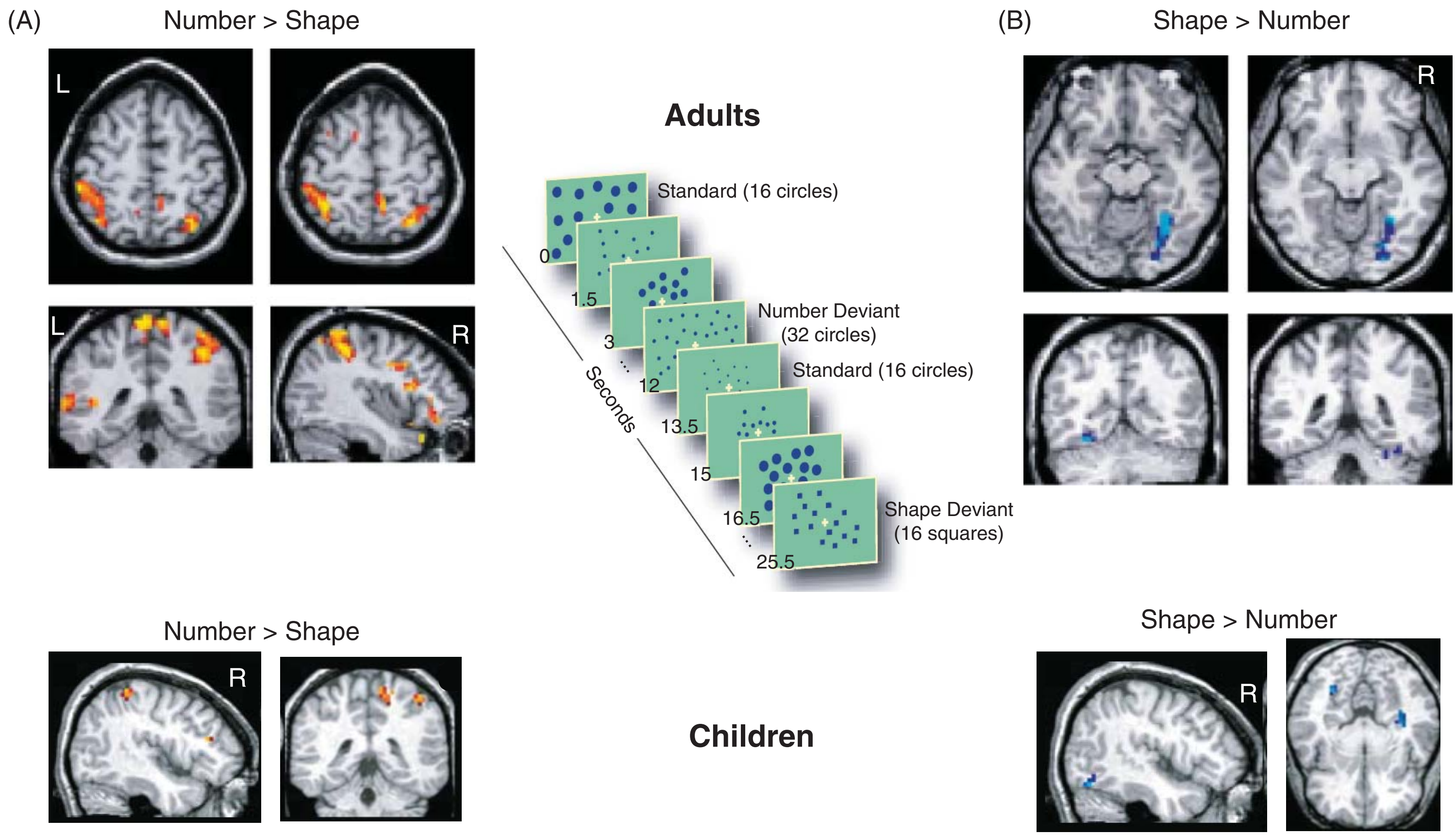
The IPS plays a crucial role in judgement of numerical quantity, and neural signatures for discrimination of small and large quantity are prominent in adults as well as children (Figure 6.3A, c–d) (Cantlon et al., 2009a; Cutini et al., 2014; Heine et al., 2011; Hyde & Spelke, 2009; Kucian et al., 2008; Piazza, Mechelli, Butterworth, & Price, 2002). The IPS is sensitive to the numerical distance between quantity represented in both discrete (Figure 6.3B, a) (Piazza et al., 2004) and analog forms (Vogel et al., 2013) (Figure 6.3B, b), and modulation of IPS activity to nonsymbolic numerosities has been reported in children as young as 4 years of age (Figure 6.3B, d) (Cantlon et al., 2006). Moreover, there is evidence to suggest fine-tuned neural discriminability of individual items in the IPS (Figure 6.3B, c) (Harvey et al., 2013). The IPS is also strongly modulated by numerical distance in symbolic representations of quantity (Ansari, Garci, Lucas, Hamon, & Dhital, 2005; Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Hinton, Dymond, Von Hecker, & Evans, 2010; Naccache & Dehaene, 2001). Taken together, these findings point to a key role of the IPS in building visual representations of quantity through neural substrates that overlap with those engaged by visual short-term memory and attention.

In contrast to the IPS, the FG plays an essential role in high-level object recognition (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Gauthier, Skudlarski, Gore, & Anderson, 2000; Goodale & Milner, 1992; Grotheer, Herrmann, & Kovacs, 2016; Holloway et al., 2013; Shum et al., 2013; Wimmer, Ludersdorfer, Richlan, & Kronbichler, 2016). Crucially, FG regions proximal to the visual word form area (McCandliss, Cohen, & Dehaene, 2003; Saygin et al., 2016) show strong responses to numerical symbols in both adults and children (Figure 6.3C) (Ansari, 2008; Cantlon et al., 2009a; Grotheer et al., 2016; Holloway et al., 2013; Shum et al., 2013). Intracranial electrophysiological recordings suggest that specific subdivisions of the FG may be differentially sensitive to number stimuli compared to perceptually similar letters and false word stimuli (Figure 6.3C, a) (Shum et al., 2013), suggesting a crucial role of the FG in perceptual decoding of numerical digits. Involvement of the FG in numerical judgments has also been supported by functional imaging studies in both adults (Figure 6.3C, b) (Holloway et al., 2013) and in children 6 to 7 years old (Figure 6.3C, c) (Cantlon et al., 2009a).

Thus, starting from the earliest years of schooling, both the dorsal and ventral visual streams, anchored in the IPS and FG, respectively, contribute to symbolic and nonsymbolic representations of quantity. Beyond this, interactions of the IPS and FG with multiple prefrontal and parietal cortical regions allow active manipulation of numerical representations over longer time periods in working memory, thereby facilitating more complex rule-based problem solving (Menon, 2014). See the section titled “Developmental Shifts from Prefrontal to Specialized Parietal Circuits.”

Multisensory Mapping: Development of STG–FG Circuits for Mapping Number Words to Symbols

Children are first introduced to symbolic representation of quantity using number words. Hence, auditory processing systems in the superior temporal gyrus (STG) and the middle temporal gyrus (MTG) may be crucial for the development of early symbolic manipulations of quantity, including counting objects and the use of number words (Carey, 2004). Neural building blocks associated with number word processing and their development have been less well studied because phonological abilities develop before a child enters kindergarten and functional brain imaging studies are only now


(A)
Number > Shape
L
L
R
Adults
Standard (16 circles)
0
1.5
3
...
12
13.5
15
16.5
...
25.5
Seconds
Number Deviant
(32 circles)
Standard (16 circles)
Shape Deviant
(16 squares)
(B)
Shape > Number
R
Number > Shape
R
Children
Shape > Number
R

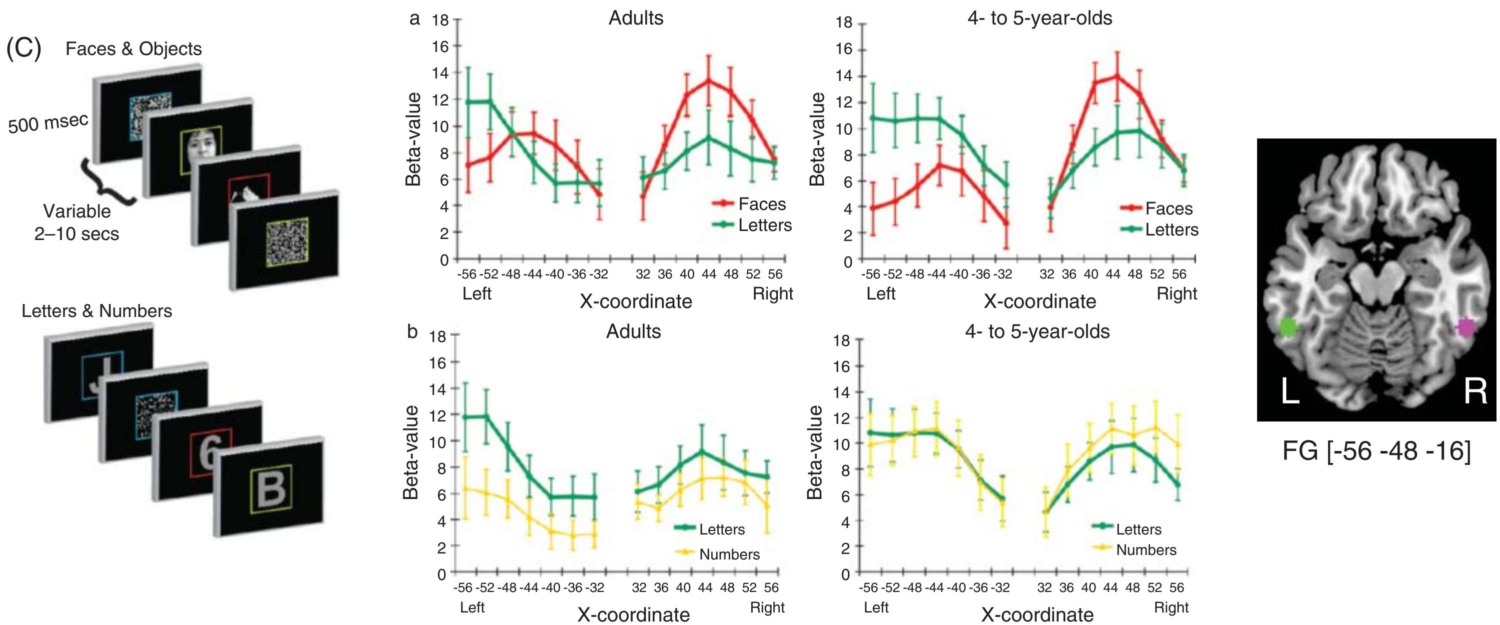


Figure 6.4 IPS and FG representations for nonsymbolic and symbolic numerosity. (A) Brain regions that are more active during the presentation of number compared to shape deviants in adults (top two rows) and 4-year-old children (bottom row). (B) Brain regions that are more active during the presentation of shape compared to number deviants in adults (top two rows) and 4-year-old children (bottom row). Adapted from Cantlon et al., 2006, © 2006 by *PLoS*. (C) Preschool children (4–5-year-olds) show adult-like differentiation in signal change in the bilateral fusiform gyrus (FG) between faces and letters, but not between letters and numbers, while adults do. Adapted from Cantlon, Pinel, Dehaene, & Pelphrey, 2011, © 2012 by the Oxford University Press. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

beginning to address the early stages of language acquisition. However, the role of the posterior STG in phonological processing is well established (Prado et al., 2011), and there is evidence to suggest that good phonological skills play a critical role in the acquisition of mathematical skills (De Smedt et al., 2010). Prior to kindergarten, children are also exposed to visual symbols, typically Arabic-Hindu digits, representing quantity.

Perceptual representation of numerical symbols is known to be supported by the VTOC, with the strongest effects in the FG (Ansari, 2008; Cantlon et al., 2006, 2011; Holloway et al., 2013; Shum et al., 2013) in both adults and children (Figure 6.3). There is evidence to suggest that perceptual processing of nonsymbolic discrete stimuli (e.g., sets of dots) and the ability to differentiate between distinct visual objects, such as faces and letters, is relatively mature by the age of 4 (Cantlon et al., 2006) (Figure 6.4A–B). However, FG sensitivity to symbolic numbers, compared to letters, is still immature at this age (Cantlon et al., 2011) (Figure 6.4C).

Integration of phonological and orthographic codes, represented in the STG and FG, respectively, is essential for successful development of word reading skills, and coactivation of printed- and spoken-language processing networks in 8-year-olds is predictive of reading scores 2 years later (Preston et al., 2016). Multisensory mapping of STG–FG circuits for mapping number words to symbols has not been investigated in the domain of mathematical cognition, but it is reasonable to assume that a similar process may apply to processing numerical symbols.

Semantic Mapping: Development of Ventral-Dorsal Stream Circuits Supports Symbolic to Nonsymbolic Mapping

With increased proficiency in representing number words and visual numbers comes automatized three-way mapping among Arabic-Hindu digits, number words, and their semantic meaning as abstract representations of quantity. This mapping arises from a dynamic interplay between visual and auditory associative cortices, anchored in the FG and STG, respectively, and their underlying numerical representations in the PPC. Investigations of this process have, however, been focused almost solely on a two-way mapping between the FG and PPC for visual number symbols (Ansari & Dhital, 2006; Cantlon et al., 2006, 2011; Holloway et al., 2013; Ischebeck et al., 2006). This circuit linking the dorsal and ventral visual stream plays an essential role in building a strong link between symbolic and nonsymbolic quantity representations. The neural substrates for representing sets of nonsymbolic quantity matures early in development (Cantlon et al., 2006; Kaufmann, Wood, Rubinsten, & Henik, 2011; Kucian et al., 2008; Kucian, Loenneker, Martin, & Von Aster, 2011; Price, Holloway, Rasanen, Vetterinen, & Ansari, 2007), but it does not reach adult-like level of precision (Piazza et al., 2004, 2010; Siegler & Opfer, 2003) until later in development (Figure 6.2E). Connectivity between the dorsal and ventral visual streams plays an essential role in children's mathematical skill development, as demonstrated recently by a study using intrinsic functional connectivity analyses, where the strength of functional coupling between FG and dorsal PPC and the prefrontal cortex PFC, significantly predicted gains in numerical abilities over development spanning an age range between 7 and 14 years of age (Evans et al., 2015) (Figure 6.5A–B–C).

Developmental Shifts from Prefrontal to Specialized Parietal Circuits

A key neurodevelopmental feature common to both symbolic and nonsymbolic magnitude judgement tasks is a "neural shift" from general processing regions in the PFC to

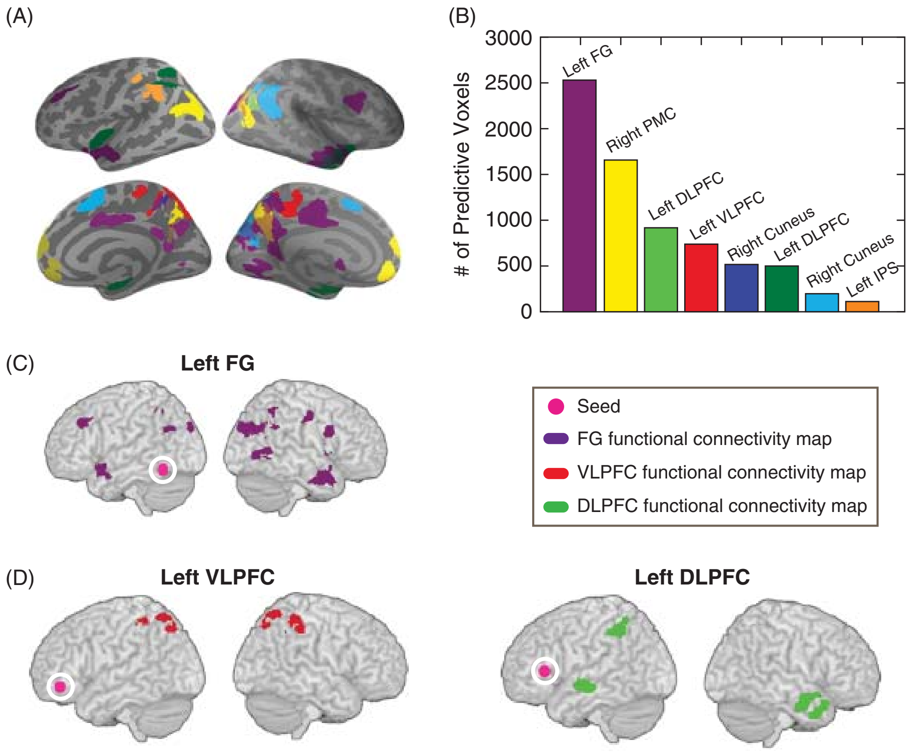


Figure 6.5 Intrinsic functional connectivity of brain regions predicts longitudinal developmental gains in numerical abilities. (A, B) FG connectivity was the most predictive of developmental gains in sample of 7- to 14-year-olds; (C) FG functional connectivity patterns that predict gains, with seed in pink, functional connectivity map in purple; (D) prefrontal cortex (PFC) functional connectivity maps that predict longitudinal developmental gain in numerical skills. Connectivity seeds in pink, functionally connectivity maps in red and green for the ventrolateral and dorsolateral PFC respectively. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
SOURCE: Adapted from Evans et al., 2015, © 2015 by SfN.

more specialized regions in the PPC. In a nonsymbolic magnitude discrimination task, Cantlon and colleagues (2009a) found that while children 6 to 7 years old engaged the bilateral inferior frontal gyrus and adjoining insular cortex, these prefrontal regions were not significantly activated in 24-year-old adults (Figure 6.6A). In contrast, both groups showed activation of the left IPS, although its spatial extent of activity was greater in adults (Figure 6.6A). Furthermore, numerical distance effects (greater activity for comparisons involving smaller ratios) have been shown to correlate with left IPS activity in adults and for a variety of manipulations, including physical size and luminance (Pinel, Dehaene, Riviere, & Lebihan, 2001). In contrast, children display the same neural effect, but in the dorsal and lateral aspect of the PFC (Figure 6.6B) (Ansari & Dhital, 2006; Ansari et al., 2005; Cantlon et al., 2009a). In a nonsymbolic comparison task, Ansari and Dhital (2006) found that only 9- to 11-year-old children displayed a distance effect in the right dorsolateral PFC, whereas 19- to 21-year-old adults had a stronger distance effect in the left anterior IPS (Figure 6.6B–C). Similarly, in a symbolic number comparison task using the

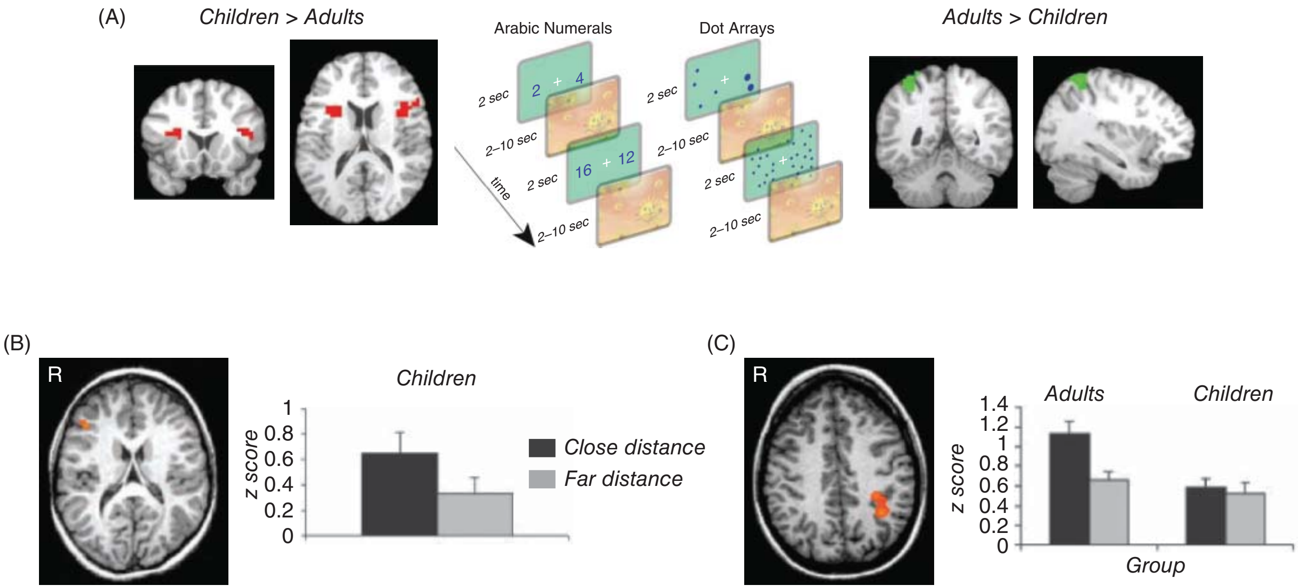


Figure 6.6 Development of IPS specialization for magnitude comparison. (A) Statistical comparison of 6-year-olds and adults during symbolic and nonsymbolic number comparison tasks. Children show greater PFC activation compared to adults while adults show greater IPS activation compared to children. Adapted from Cantlon et al., 2009, © 2009 by the Massachusetts Institute of Technology. (B) Significant distance effect during a nonsymbolic comparison task in 9-year-old children in the dorsolateral PFC (DLPFC). (C) The left IPS shows a significant distance effect in adults but not in 9-year-old children. Adapted from Ansari and Dhital, 2006, © 2006 by the Massachusetts Institute of Technology. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

same age groups, Ansari et al. (2005) found that adults showed sensitivity to numerical distance bilaterally in the IPS, whereas children showed sensitivity to numerical distance in the right precentral gyrus and right inferior frontal gyrus. Thus, although the precise neural locus of development varies by cognitive process and stimulus, a consistent profile of decreased reliance on the PFC and increased reliance on the PPC has been found as a function of development in a wide range of studies of numerical cognition involving magnitude estimation with symbolic as well as nonsymbolic stimuli.

NEURODEVELOPMENTAL PROCESSES IN MATHEMATICAL REASONING

Parietal-Frontal Working Memory Systems in the Development of Mathematical Cognition

The first hint of the possible involvement of parietal-frontal circuits in mathematical cognition came from a study of a 59-year-old patient with an extended left parietal-frontal infarct who had selective deficits in performing numerical computations with numbers greater than 4 (Cipolotti, Butterworth, & Denes, 1991). Early neuroimaging studies provided stronger evidence for the specific involvement of a core parietal-frontal working memory network in mathematical reasoning based on overlapping PPC and PFC activations in tasks involving numerical problem solving and working memory (Grabner et al., 2009; Ischebeck, Zamarian, Egger, Schocke, & Delazer, 2007; Rivera et al., 2005). Crucially, subsequent studies have shown that the role of the parietal–frontal working memory network in numerical problem solving depends on problem complexity (Ischebeck et al., 2006), individual differences in task proficiency (Dumontheil & Klingberg, 2012; Metcalfe, Ashkenazi, Rosenberg-Lee, & Menon, 2013), and developmental stage of skill acquisition (Menon, 2014).

The engagement of the parietal–frontal system changes dramatically with development (Menon, 2014). For example, Rivera and colleagues (2005) found that relative to adults, children tend to engage the PPC less and the PFC more when solving arithmetic problems, likely reflecting the increased role of visuospatial processing and the concurrent decrease in demands on cognitive control (Figure 6.7). Extending this finding, analysis of neurodevelopmental profiles among children, adolescents, and young adults revealed that differential recruitment of specific brain systems in adolescence underlies adult levels of problem-solving skills (Chang, Metcalfe, Padmanabhan, Chen, & Menon, 2016). While activation in bilateral ventral IPS increased linearly with age (Figure 6.8A), the left anterior supramarginal gyrus (SMG) subdivision (SMG-PF) (Caspers et al., 2006) showed an inverted U-shaped profile across age groups such that adolescents exhibited greater activation than both children and young adults (Figure 6.8B). Critically, greater SMG–PF activation was correlated with task performance only in adolescents (Figure 6.8B); adolescents also showed greater task-related functional connectivity of the SMG–PF with the ventro-temporal, anterior temporal, and prefrontal cortices relative to both children and adults. These results suggest that nonlinear up-regulation of the SMG–PF and its interconnected functional circuits facilitate adult-level performance in adolescents, and crucially, they demonstrate how anatomically precise analysis of both linear and nonlinear neurofunctional changes with age is necessary for more fully characterizing cognitive development.

Other studies have addressed the link between working memory abilities and

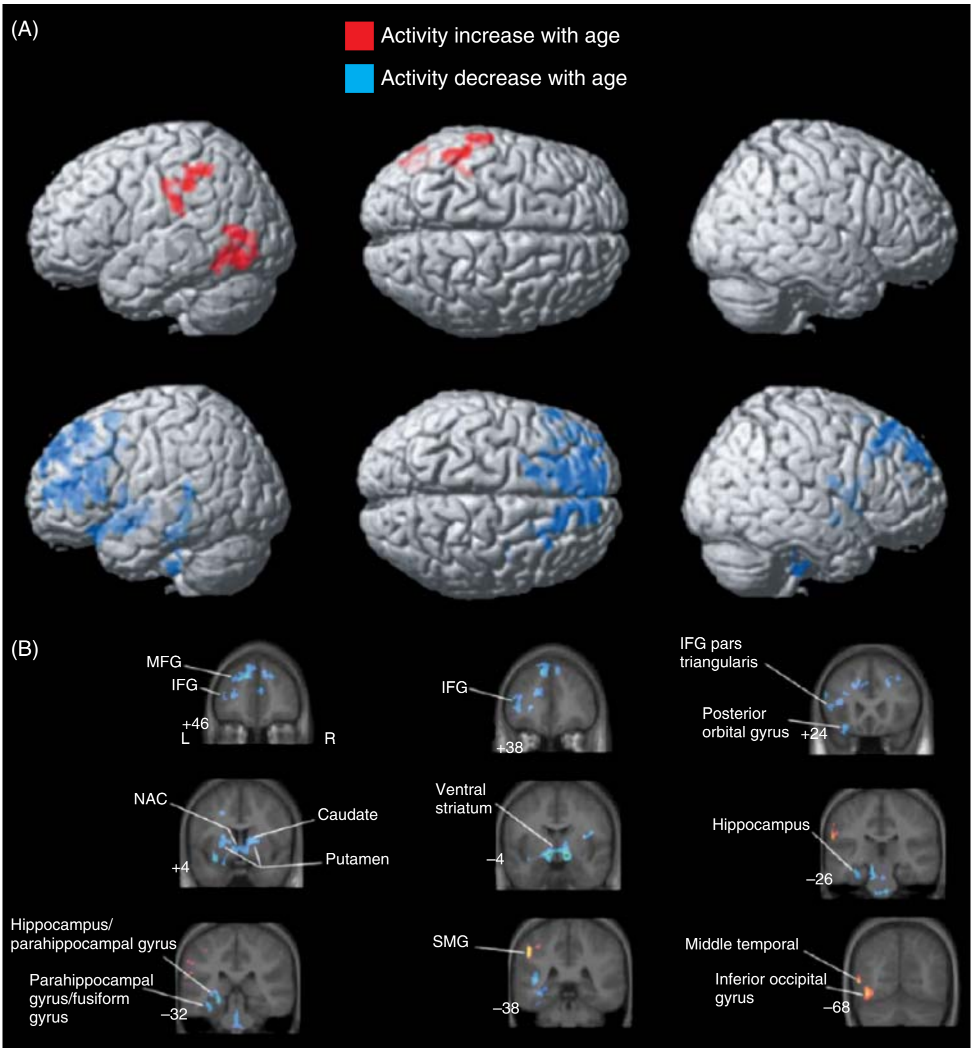


Figure 6.7 Fronto-parietal shift in activation over development. (A) Lateral view: Brain areas in which activity during arithmetic problem solving increases or decreases with age. Increases were detected in the supramarginal gyrus (SMG) and the ventral temporal occipital cortex (VTOC), while decreases were detected in the lateral prefrontal cortex as well as in the superior and middle temporal gyri. (B) Coronal view: Brain areas that show activity decreases with age also included the medial temporal lobe (MTL), in the hippocampus and parahippocampal gyrus. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
Source: Adapted from Rivera et al., 2005, © 2005 by the Oxford University Press.

mathematical problem-solving skills more directly. Dumontheil and Klingberg (2012) found that IPS activity during a visuospatial working memory task predicted arithmetic performance 2 years later in a sample of 6- to 16-year-olds. Further analyses on the neural correlates of individual components of working memory have provided evidence for the fractionation of neurofunctional systems associated with distinct working memory components during arithmetic problem solving (Arsalidou & Taylor, 2011; Metcalfe

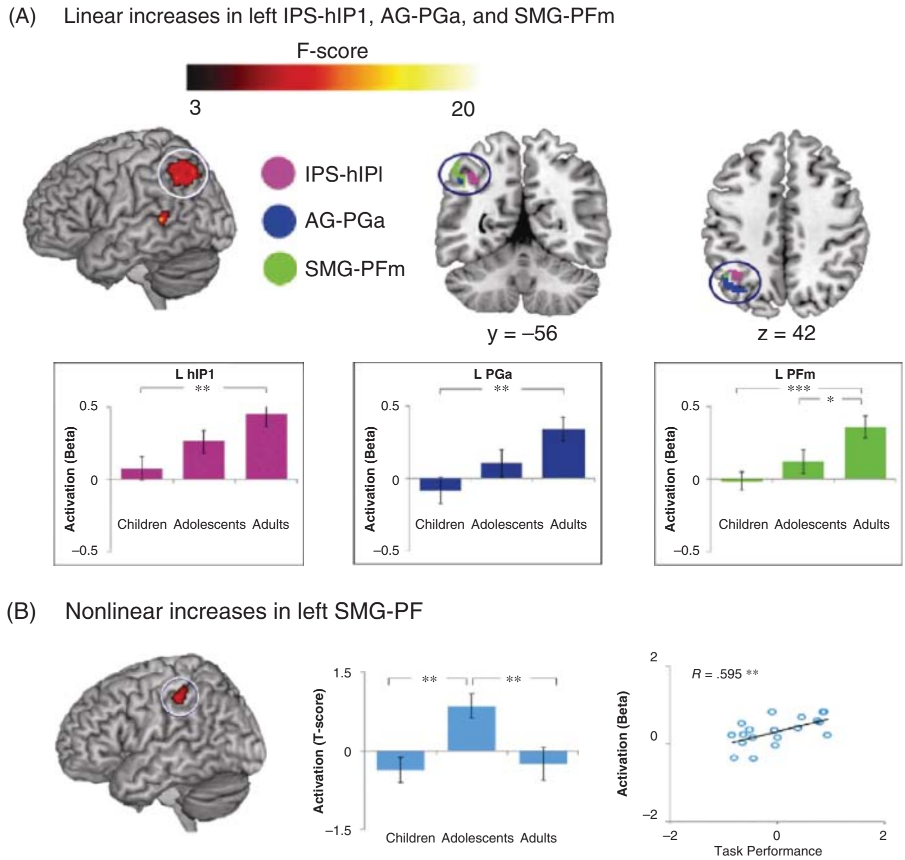


Figure 6.8 Developmental changes in parietal cortex from childhood to adolescence and adulthood. (A) Linear increases with age. Cytoarchitectonic probabilistic labeling indicates that activations during arithmetic problem solving in both ventral lateral subdivisions of the IPS (hIP1) and adjoining anterior angular gyrus–PGa (AG–PGa), and posterior subdivision of the supramarginal gyrus (SMG–PFm) show linear increases from childhood to adulthood. (B) Nonlinear increases with age. Activation in the left anterior supramarginal gyrus (SMG) showed an inverted U-shaped profile across age groups. Greater SMG–PF activation was correlated with task performance in adolescents only. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
SOURCE: Adapted from Chang, Rosenberg-Lee, Metcalfe, Chen, & Menon, 2015, © 2015 by Elsevier.

et al., 2013; Rottschy et al., 2012). Specifically, analysis of the relation between the central executive, phonological, and visuospatial components of working memory and brain activation during an arithmetic verification task in a large group of 7- to 9-year-old children revealed that visuospatial working memory is a strong predictor of mathematical ability in children in this age group. This study also demonstrated that visual working memory is associated with increased problem complexity-related responses in left dorsolateral and right ventrolateral PFC as well as in the bilateral IPS and SMG. Visuospatial working memory and the central executive component were associated with largely distinct patterns of brain responses during arithmetic problem solving, and overlap was observed only in the ventral aspects of the left SMG (Figure 6.9), suggesting that this region is an important locus for the integration of information in working

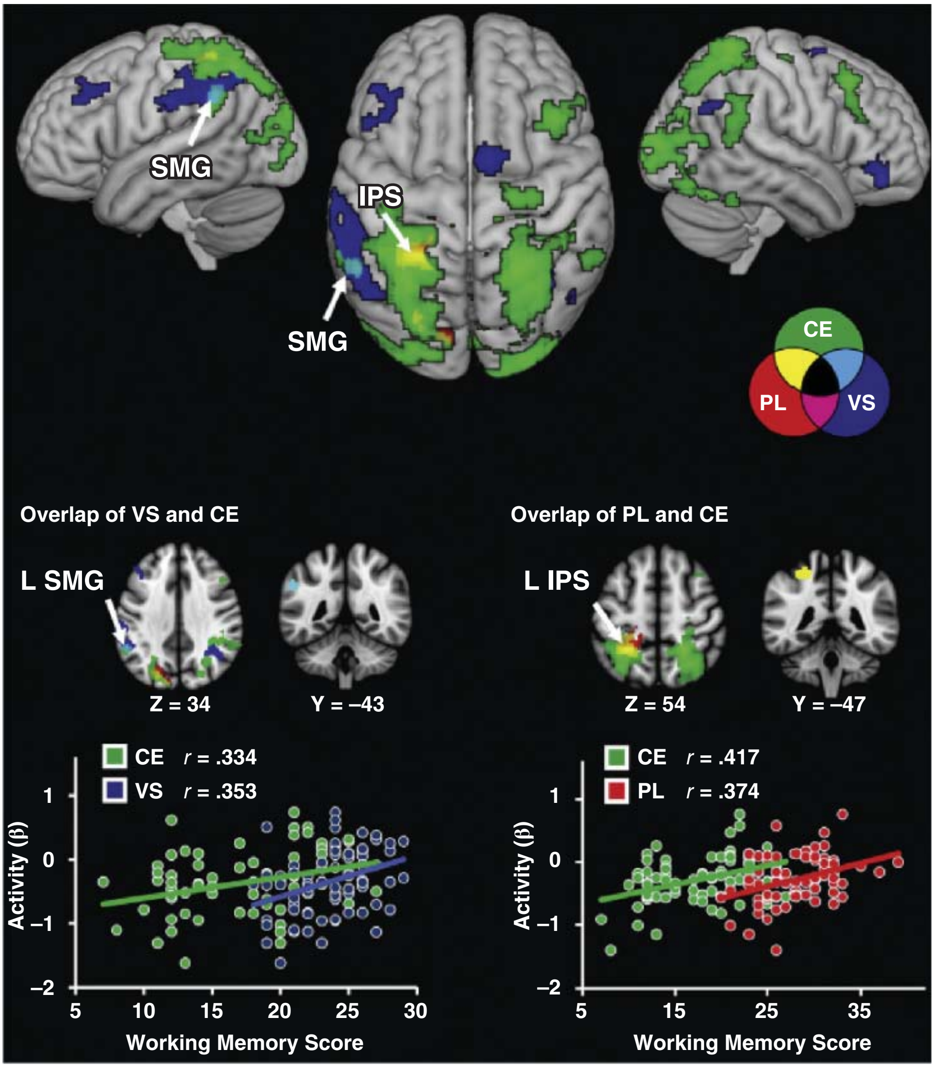


Figure 6.9 Functional dissociations and overlap between brain areas associated with working memory components in 7- to 9-year-old children during arithmetic problem solving. Top panel: Overlap between the central executive (CE) and visuo-spatial (VS) components was observed only in the left supramarginal gyrus (SMG); overlap between CE and phonological loop (PL) components was observed only in the left intraparietal sulcus (IPS); no overlap was observed between VS and PL components. Negative correlation between activity and PL ability is not depicted. No overlap for VS and PL (magenta) was observed. Bottom panel: Coronal slices depict regions of interest selected as overlap in correlations of activity and individual working memory components. Scatter plots are based on functional clusters identified using whole-brain regression analysis and are provided for the purpose of visualization. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
Source: Adapted from Metcalfe et al., 2013, © 2013 by John Wiley & Sons.

memory during arithmetic problem solving in children (Ansari, 2008; Dehaene et al., 2003; Kawashima et al., 2004; Kucian et al., 2008; Menon et al., 2000; Rivera, Reiss, Eckert, & Menon, 2005; Rosenberg-Lee et al., 2014). Finally, analysis of intrinsic functional connectivity suggests that a net-work of prefrontal and parietal cortical areas supports the longitudinal development of numerical abilities from 7 to 14 years of age (Evans et al., 2015) (Figure 6.5D). These findings further confirm the pivotal role of overlapping parietal–frontal working memory circuits in children's mathematical skill development.

Development of Unique Representations for Distinct Numerical Problems

Although localization of brain activation has provided useful knowledge about the relative engagement of task-specific brain areas during problem solving, it offers limited insights into the cognitive and neural representations of distinct numerical problems in the brain. Even in brain areas where the level of functional engagement is mature by age 10, the underlying neural representations continue to be refined. Studies examining fine-grained representations using multivariate pattern analysis (Kriegeskorte & Kievit, 2013) demonstrated that the maturation of arithmetic problem-solving skills is characterized by the emergence of common neural representations across distinct arithmetic problem types. For instance, addition and subtraction are two complementary arithmetic operations that manipulate the same quantity in opposite directions. In most elementary school curricula, addition is learned as the primary operation, whereas subtraction builds on prior knowledge of addition and the inverse relationship of subtraction with addition (Campbell, 2008). In a recent study, Chang and colleagues (2015) showed that, unlike adults, children do not show invariant neural representations for distinct problem types in the IPS, FG, PFC, and anterior temporal cortex (ATC). This finding suggests that the emergence of arithmetic problem-solving skills from childhood to adulthood is characterized by maturation of common neural representations between distinct numerical operations and involves distributed brain regions important for representing and manipulating numerical quantity (Figure 6.10).

Medial Temporal Lobe: System for Associative Learning

Over the past few years, evidence has been accumulating for the differential involvement of the medial temporal lobe (MTL) declarative memory system in mathematical learning, especially during key stages of the development of cognitive skills (Qin et al., 2014) (Figure 6.11 A–B–C). The importance of the MTL, particularly its hippocampal subdivision, in learning and memory consolidation for events in space and time is well known (Davachi, 2006; Davachi, Mitchell, & Wagner, 2003; Diana, Yonelinas, & Ranganath, 2007; Eichenbaum, Yonelinas, & Ranganath, 2007; Squire, 1992; Squire, Genzel, Wixted, & Morris, 2015; Squire, Stark, & Clark, 2004; Tulving, 1983). Theories of memory consolidation posit that the hippocampus plays an important role in the early stages of learning and retrieval, but its involvement decreases over time with concomitant increases in the reliance on neocortical memory systems. Thus, despite its critical role in learning and memory formation, hippocampal contributions to mathematics learning and cognitive development more broadly have received little attention until recently.

The first evidence for the differential engagement of the hippocampal memory system in mathematical cognition over development came from a cross-sectional study in children, adolescents, and adults spanning the ages between 8 and 19 (Rivera et al., 2005). Children exhibited significantly greater engagement of multiple MTL regions, including the hippocampus (Figure 6.7). Similarly, De Smedt and colleagues (2011) found greater hippocampal response in children compared to adults when solving addition problems, although hippocampal activation was not detected for subtraction

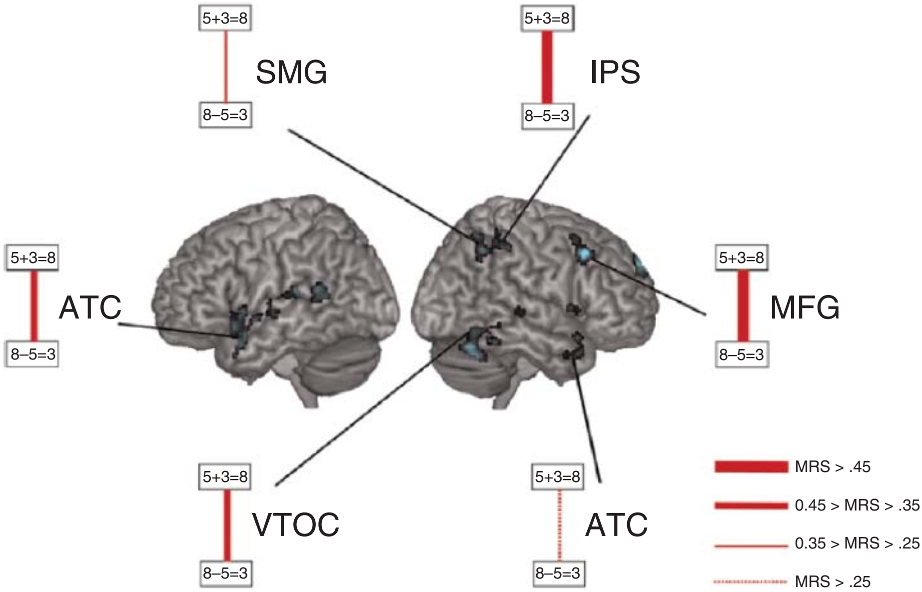


Figure 6.10 Neurodevelopmental changes in multivariate representational similarities between distinct numerical problems. Multivariate representational similarity (MRS) was used to assess similarity between addition and subtraction problems. The IPS and middle frontal gyrus (MFG) showed the strongest effects in developmental changes, followed by the anterior temporal cortex (ATC), ventral-temporal occipital cortex (VTOC), and supramarginal gyrus (SMG). Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
SOURCE: Adapted from Chang et al., 2015, © 2015 by Elsevier.

problems. This finding could be due to the fact that subtraction problems are less well rehearsed and more difficult to memorize, mainly because subtraction problems are not commutative. These findings highlight the dynamic role of the hippocampus in the maturation of memory-based problem-solving strategies and its greater engagement in childhood followed by decreased involvement in adolescence and adulthood (Figure 6.11B).

As noted, children's gains in problem-solving skills during early school years are characterized by the gradual replacement of inefficient procedural strategies (i.e., counting) with direct retrieval of math-relevant facts (Figure 6.11A) (Cho, Ryali, Geary, & Menon, 2011; Geary, 2011; Geary & Brown, 1991; Geary & Hoard, 2003; Qin et al., 2014). Cho and colleagues (2012) examined neurodevelopmental changes related to increased use of retrieval strategies and found that higher retrieval fluency was associated with greater response in multiple brain regions, including the hippocampus and parahippocampal gyrus subdivisions of the MTL. Thus, children's use of retrieval strategies, far from being idiosyncratic, is in fact associated with a predictable profile of hippocampal responses. A related study found that retrieval and counting strategies were associated with different activation patterns in hippocampal regions important for memory encoding and retrieval, including the bilateral hippocampus and parahippocampal gyrus (Cho et al., 2011). The existence of decodable fine-scale pattern differences in the spatial pattern of functional magnetic resonance imaging (fMRI) signals suggests not only that the hippocampus is differentially

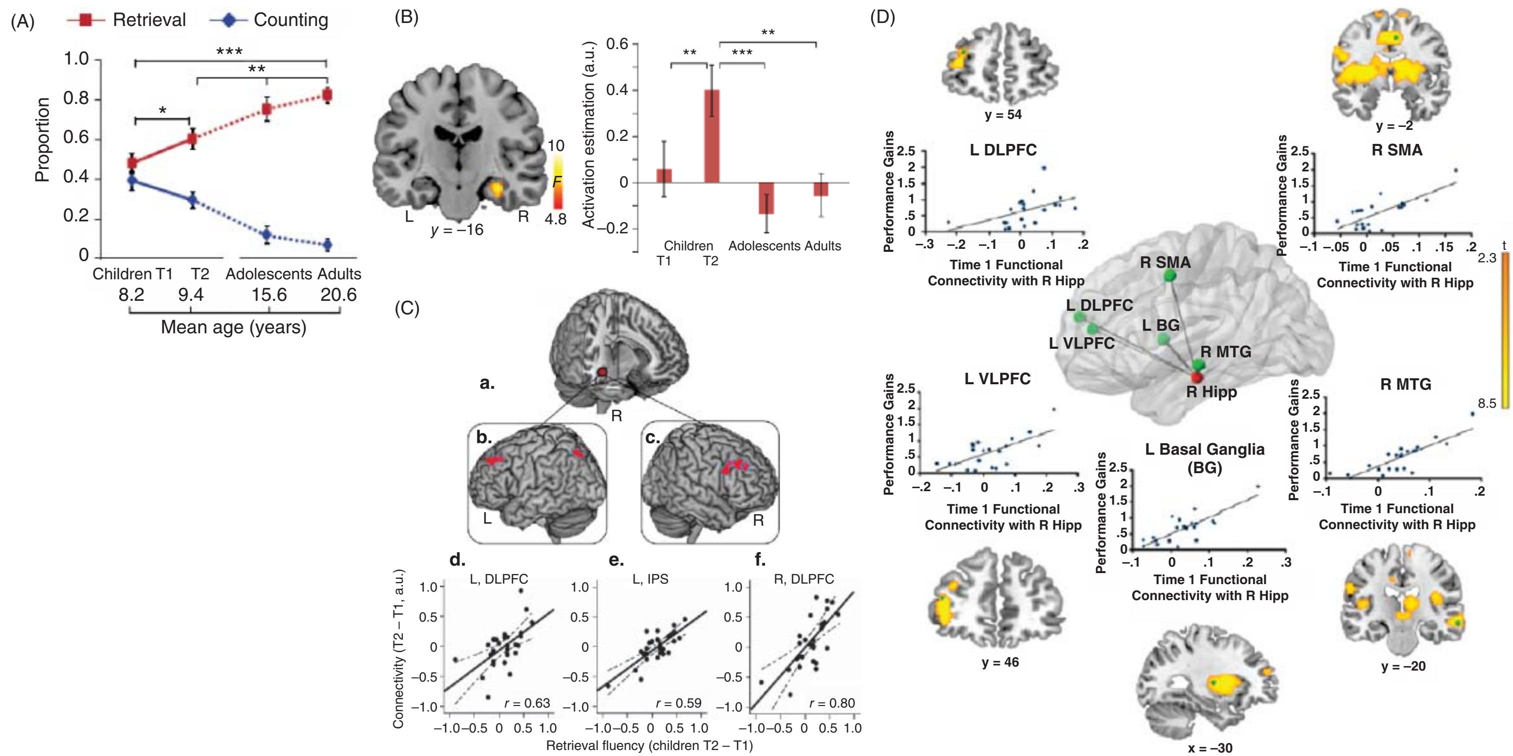


Figure 6.11 Maturation of medial temporal lobe (MTL) engagement and connectivity. (A) Longitudinal changes in strategy use during arithmetic problem solving from childhood to adulthood. Gradual increases in memory-based strategies and decreases in counting strategies. (B) Right hippocampal response showing main effect of group across children at different time points (Time 1 and Time 2 approximately 1.2 years apart), adolescents, and adults. Bar graphs depict developmental changes in the functionally defined hippocampus cluster. (C) Longitudinal changes in hippocampal–neocortical functional circuits in relation to individual improvements in children's use of memory-based problem-solving strategies; (a) Right hippocampus seed region used in effective connectivity analysis; (b, c) Left and right dorsolateral prefrontal cortex (DLPFC) and the left intraparietal sulcus (IPS) regions that showed increased effective connectivity with the hippocampus, as a function of longitudinal improvements in retrieval fluency from Time 1 to Time 2; (d–f) Scatter plots depict the relation between longitudinal changes in retrieval fluency (*x*-axes) and changes in effective connectivity strength from Time 1 to Time 2 (*y*-axes). Adapted from Qin et al., 2014, © 2014 by NPG. (D) Intrinsic connectivity of the hippocampus predicts improvement in numerical abilities in 7- to 9-year-old children following 2 months of cognitive intervention. Performance gains were correlated with Time 1 hippocampal connectivity with the left DLPFC, left ventrolateral prefrontal cortex (VLPFC), right supplementary motor area (SMA), left basal ganglia (BG), and right middle temporal gyrus (MTG). Adapted from Supekar et al., 2013, © 2013 by National Academy of Sciences. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

engaged in relation to retrieval but also that the underlying neural resources are accessed and used differently during each strategy. In a seminal study, Qin and colleagues (2014) demonstrated that children's transition from counting to memory-based retrieval strategies (Figure 6.11A) over a 1.2-year interval was mediated by increased hippocampal activation (Figure 6.11B) and increased hippocampal–neocortical connectivity (Figure 6.11C). Following an initial increase in hippocampal engagement during middle childhood, this hippocampal dependency decreased during adolescence and adulthood (Figure 6.11B) despite further improvements in memory-based problem solving (Figure 6.11A). This pattern of initial increase and subsequent decrease in activation provides novel support for models of long-term memory consolidation, which posit that the hippocampus plays a time-limited role in the early phase of knowledge acquisition (McClelland, McNaughton, & O'Reilly, 1995; Tse et al., 2007). Consistent with this pattern of developmental change, previous studies of adults have reported no reliable hippocampal engagement during basic arithmetic tasks. Thus, the hippocampal system is critical to children's early learning of arithmetic facts (Cho et al., 2011, 2012; De Smedt, Holloway, & Ansari, 2011), while retrieval is largely dependent on the neocortex in adults (Dehaene et al., 2003; Menon, 2014).

The shift from procedural to memory-based retrieval strategies and increased hippocampal activation is accompanied by decreases in parietal–frontal engagement (Qin et al., 2014). Coincidentally, increases in functional connectivity between hippocampal and neocortical circuits (Figure 6.11C) were significantly related to longitudinal improvements in retrieval fluency. In a tutoring study designed to facilitate rapid retrieval of math facts, hippocampal–PFC functional circuits predicted performance gains over an eight-week interval (Figure 6.11D**).** Specifically, children who exhibited higher intrinsic functional connectivity in these circuits prior to tutoring showed the greatest performance improvement in math problem solving (Supekar et al., 2013). Together, these findings suggest that hippocampal–neocortical circuit reorganization therefore plays an important role in children's shift from effortful counting to more efficient memory-based problem solving.

Cognitive Control Systems in Mathematical Cognition

Prefrontal control processes are important for virtually every complex cognitive task, including mathematical reasoning (Shallice & Evans, 1978). The role of both working memory and declarative memory systems in mathematical reasoning therefore must be considered in the context of cognitive control processes that support flexible problem solving and learning. Prefrontal control processes serve various functions in numerical cognition, including maintenance of attention on goal-relevant numerical representations, online manipulation of information in working memory, inhibition of irrelevant information, and implementation of task-relevant processes. Implementation of such control relies on dynamic functional interactions among multiple frontal regions (Cai, Ryali, Chen, Li, & Menon, 2014; Cai et al., 2016; Cole et al., 2013; Ham, Leff, De Boissezon, Joffe, & Sharp, 2013; Seeley et al., 2007; Sridharan, Levitin, & Menon, 2008), and recent research has begun to elucidate the role of parietal–frontal and hippocampal–frontal circuits in different aspects of cognitive control during mathematical cognition.

As previously noted, mathematical problem solving is often supported by the co-engagement of parietal and prefrontal regions associated with working memory functions. Children as young as 7 show reliable, and consistent, patterns of brain activity during arithmetic problem solving in multiple PFC regions (Houde, Rossi, Lubin, & Loliot, 2010). Commonly activated PFC regions include the anterior insula and the ventrolateral and dorsolateral PFC (VLPFC and DLPFC) (Houde et al., 2010), and these regions are also implicated in a wide range of cognitive control tasks in adults as well as children. This profile of anatomical overlap suggests a common mechanism by which maturation of basic cognitive control systems can influence skill development across multiple cognitive domains, including mathematical reasoning skills.

Efficient control requires the concerted coordination among multiple brain regions, and there is growing evidence to suggest that such coordination is implemented via dedicated neurocognitive networks. Two key networks play a fundamental role in cognitive control processes in the human brain: The insula–cingulate salience network (SN), which includes the anterior insula and anterior cingulate cortex (ACC), and the dorsal parietal–frontal working memory attentional network, which includes the VLPFC and DLPFC and SMG (Figure 6.1). The anterior insula in the salience network has been shown to be a major causal hub initiating control signals during arithmetic problem solving, and its role is progressively refined over development (Supekar & Menon, 2012) (Figure 6.12A). In particular, the causal influence of the anterior insula over the VLPFC and DLPFC and ACC is significantly greater in adults, compared to 7-year-old children (Figure 6.12B). Moreover, despite higher levels of PFC activation in children, the strength of their causal modulatory influences (Figure 6.12B) and their structural connectivity (Figure 6.12C) to the parietal cortex has been shown to be significantly weaker relative to adults.

Crucially, weaker PFC control signals have been shown to be associated with lower levels of arithmetic performance, and network interactions better predicted reaction time in both children and adults. In children, the strength of casual signals from the anterior insula to the SMG and VLPFC significantly predicted reaction times. Conversely, in adults, the strength of the functional coupling between the anterior insula and the SMG, VLPFC, and ACC predicted reaction times. It is noteworthy that even though a different set of links predicted reaction times in both groups, the anterior insula–SMG link was common in both cohorts. Thus, multiple PFC control signals contribute to efficient problem-solving skills in adults, and weak signaling mechanisms contribute to lower levels of performance in children.

Dynamic causal modeling of fMRI data has provided further insights into the temporal profile of interactions among brain regions involved in mediating retrieval fluency during arithmetic problem solving. These regions include both the ventrolateral and the dorsolateral aspects of the PFC, and their interaction with MTL systems (Figure 6.13A–B–C) (Cho et al., 2012). Particularly, causal analysis has revealed strong bidirectional interactions between the hippocampus and the left VLPFC and DLPFC during math problem solving (Figure 6.13D). Crucially, causal influences from the left VLPFC to the hippocampus may act as the main "top-down" component, while causal influences from the hippocampus to the left DLPFC might serve as the main "bottom-up" component of this retrieval network. Although they are still preliminary, these analyses highlight the differential contribution of hippocampal–prefrontal circuits to the early development of retrieval fluency in arithmetic problem solving and provide a

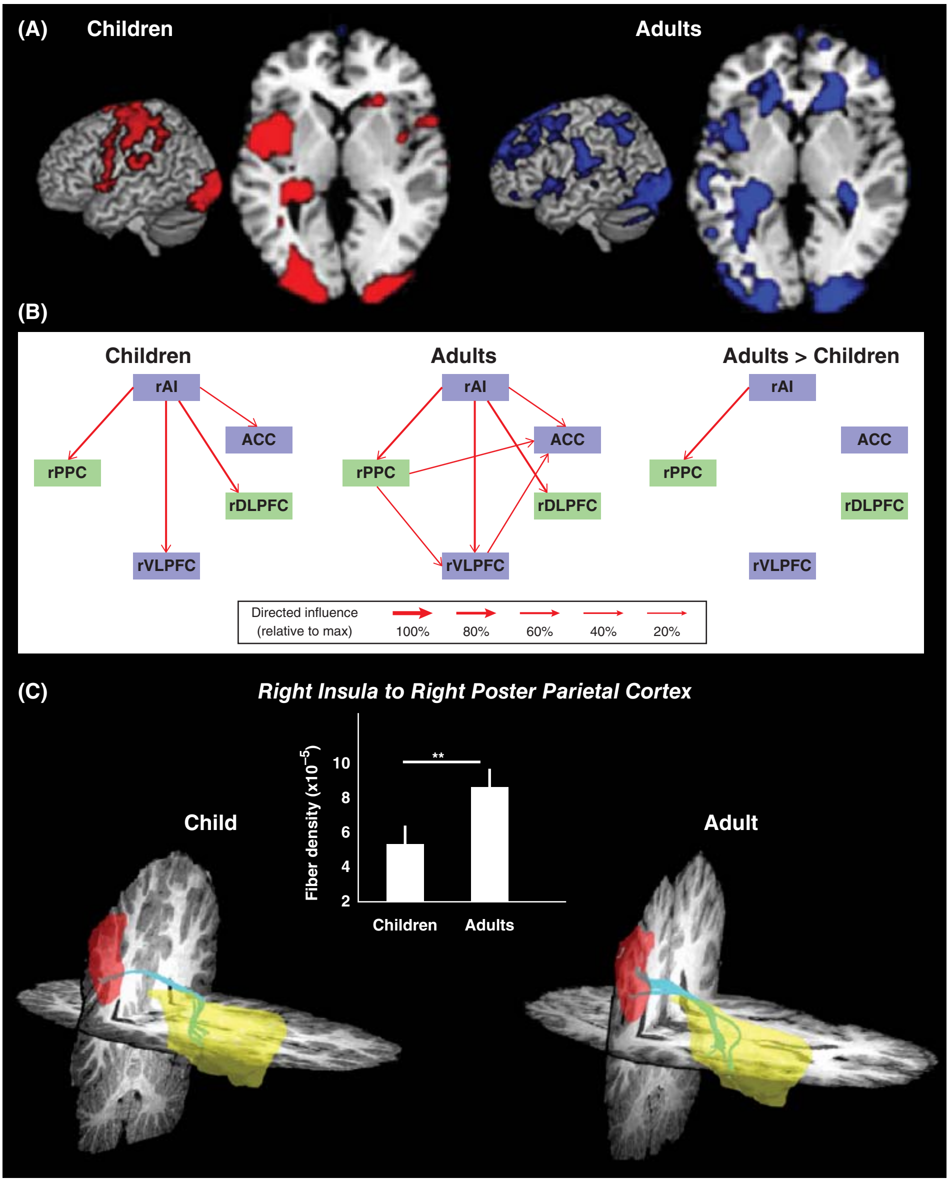


Figure 6.12 Developmental changes in cognitive control networks associated with numerical problem solving. (A) Brain activation in the salience network (SN) and the dorsal parietal–frontal working memory attentional network during arithmetic problem solving in children (left) and adults (right). (B) Weaker dynamic causal interactions between the SN and the working memory attentional network in children compared to adults. (C) Developmental changes in white matter tracts linking the SN and the working memory attentional network. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.
SOURCE: Adapted from Supekar & Menon, 2012, © 2012 by *PLoS*.

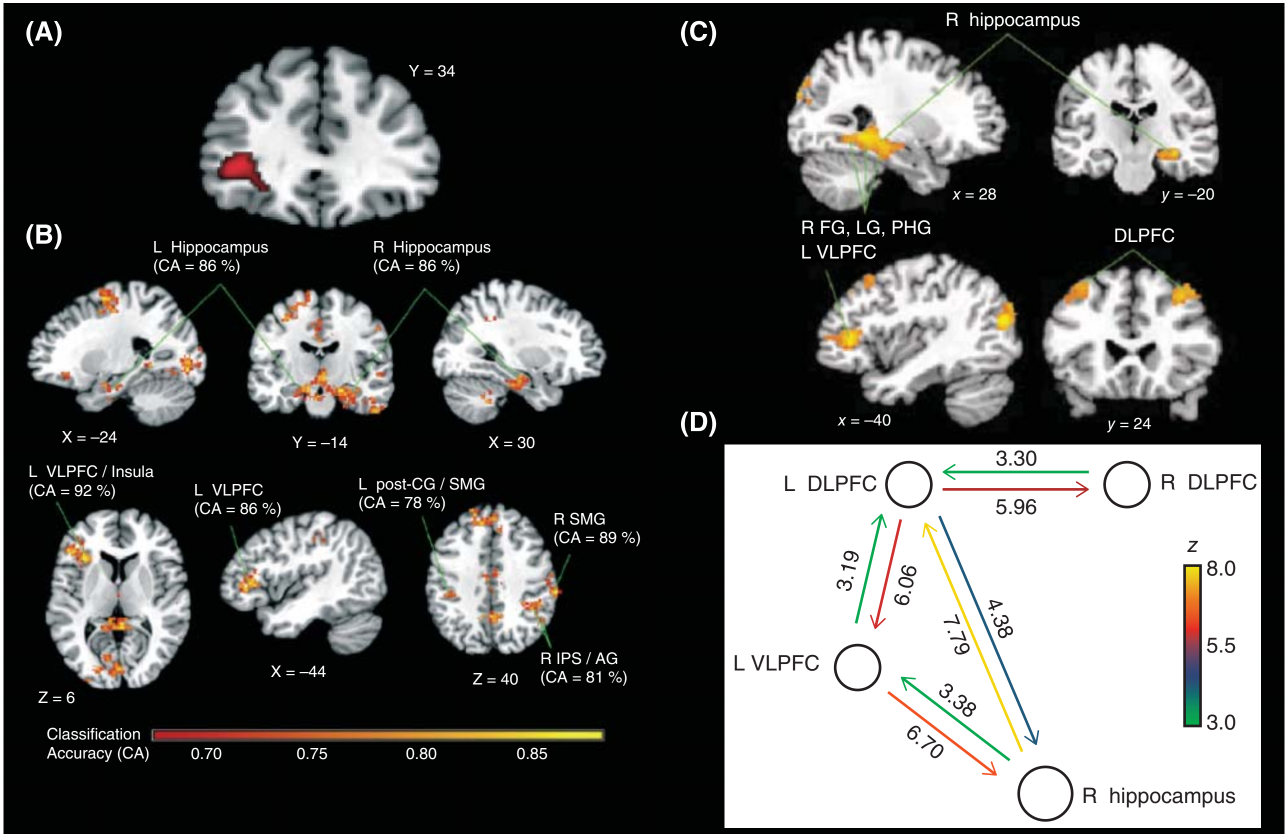


Figure 6.13 Hippocampal-prefrontal retrieval network. (A) The left ventrolateral PFC (VLPFC) showed greater activation in 7- to 9-year olds who used retrieval strategies (retrievers), compared to those who primarily used counting strategies (counters). (B) Multivariate pattern analyses revealed significant differences in spatial activation patterns between retrievers and counters in the bilateral hippocampus and adjoining parahippocampal gyrus, in the left VLPFC and adjoining anterior insula, as well as in the posterior parietal cortex (PPC) in the bilateral supramarginal gyrus (SMG), right intraparietal sulcus (IPS), and adjacent angular gyrus (AG). Peak classification accuracies are shown in parentheses. Adapted from Cho et al., 2011, © 2015 by John Wiley & Sons. (C) Brain responses associated with greater retrieval fluency. Activity levels were positively correlated with retrieval fluency in the right hippocampus and adjacent parahippocampal gyrus (PHG) as well as fusiform and lingual gyri (FG and LG), left VLPFC, and bilateral DLPFC. (D) Multivariate dynamical systems modeling of hippocampal–prefrontal retrieval network. The left VLPFC and the left DLPFC showed highly direct causal influences with the right hippocampus. Adapted from Cho et al., 2012, © 2012 by the Massachusetts Institute of Technology. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

novel framework for studying dynamic developmental processes involving the hippocampus and PFC that accompany the maturation of mathematical skills. Further research is needed to investigate how these processes contribute to concomitant improvements in cognitive control over retrieval, including successful inhibition of irrelevant information, such as incorrect answers, intermediate steps, and operand intrusions (Barrouillet & Lepine, 2005; Passolunghi & Siegel, 2004), which are all cognitive functions known to be mediated by VLPFC regions.

CONCLUSIONS AND FUTURE DIRECTIONS

The ability to represent and manipulate numbers is thought to be innate and phylogenetically determined. An evolutionary conserved system for numbers (i.e., the "number sense"), which humans share with primates as well as with less evolved vertebrate and invertebrate species, has been proposed to set the foundations of mathematical reasoning. In humans, this system is anchored in the PPC, a heterogeneous brain region that supports multiple visuospatial cognitive functions, including visuospatial attention and short-term memory. Perceptual and semantic representation of quantity in the VTOC and PPC supports the development of mathematical reasoning via multiple functional brain circuits. The core building blocks of mathematical reasoning develop through a hierarchical representation of neurocognitive functions comprising visual and auditory association cortices, which decode the visual and phonological features of numerical stimuli. A parietal-attentional system then helps build amodal semantic representations of numerosity by combining perceptual inputs with visuo-spatial cognitive primitives.

Converging evidence from studies of infants (Feigenson et al., 2004), preschool children (Cantlon et al., 2006), and adults (Ansari, 2008) as well as nonhuman primates (Cantlon & Brannon, 2006; Nieder, Freedman, & Miller, 2002), has revealed that the representation of approximate quantities is supported by the IPS in the dorsal aspects of the PPC. In addition to the dorsal PPC, the VTOC also plays an important, though often underappreciated, role in number processing. The brain builds arithmetic skills with the support of these systems, but this is only one part of the necessary circuitry. Mathematical reasoning relies on and requires multiple cognitive systems involving working memory, episodic and semantic memory systems, and executive control functions. Both developmental studies in young children and training studies in adults are beginning to highlight the important role of these systems in building new representations in the dorsal PPC and in the VTOC. Furthermore, as task manipulations become more sophisticated with better matching of control tasks on multiple dimensions, multivariate approaches will likely be more useful for examining distinct neural representations in these brain areas.

There is now growing evidence to suggest that functional circuits engaged by children are not the same as those engaged by adults, who have evolved multiple strategies and schema-like knowledge for efficient learning. Hence, a number of scaffolding systems are likely to be engaged during development to support the efficient acquisition of mathematical reasoning skills. Critically, studies are now beginning to investigate the role of parietal–frontal and hippocampal–frontal circuits that otherwise might be overlooked in studies involving only adults. These data converge on the idea that the precise nature of this engagement is a function of developmental stage, domain knowledge, problem complexity, and individual proficiency in use of efficient problem-solving strategies.

We are still in the initial stages of understanding how functional brain circuits unfold with development. It is nevertheless clear that the exclusive focus on activity levels in a small set of brain regions identified in highly skilled adults will likely miss important changes in network-level functional organization that accompany learning and development associated with schooling. Increasingly, the focus has also shifted to multivariate analyses, as it is evident that similar levels of activation across task conditions do not necessarily imply similar kinds of information processing (Blair, Rosenberg-Lee, Tsang, Schwartz, & Menon, 2012; Prado et al., 2011; Raizada et al., 2010). These types of fine-grained analyses clearly have important implications for understanding brain mechanisms mediating the formation of unique stimulus representations and how they mature with learning and development (Ashkenazi, Rosenberg-Lee, Tenison, & Menon, 2012; Chang et al., 2015).

Most previous normative adult and developmental studies of mathematical reasoning have focused mainly on localization of activation and age-related changes, but it is becoming increasingly clear that cognition depends on interactions within and between large-scale brain networks (Bressler & Menon, 2010). New research is beginning to highlight the significant and specific changes in frontal to posterior functional connectivity that take place during a time period important for mathematical skill development. In the long run, a systems neuroscience approach, with its emphasis on networks and connectivity, rather than a pure localization approach, is better suited to further understanding how complex skills, such as arithmetic, develop and are expressed in adulthood. Mathematical reasoning requires the integration of multiple cognitive processes, which rely on the engagement of distributed brain areas subserved by long-range connections that undergo significant changes with development (Fair et al., 2008; Supekar & Menon, 2012; Supekar et al., 2010). The recruitment of specific brain circuits changes dynamically as a function of training and development. In the many different ways we have highlighted in this chapter, mathematical reasoning serves as a model domain for investigating the ontogenesis of human cognitive and problem-solving skills.

REFERENCES

Agrillo, C., Dadda, M., & Bisazza, A. (2007). Quantity discrimination in female mosquitofish. *Animal Cognition*, *10*, 63–70.

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Animal Cognition*, *11*, 495–503.

Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex*, *4*, 544–554.

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *National Review of Neuroscience*, *9*, 278–291.

Ansari, D., & Dhital, B. (2006). Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: an event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *18*, 1820–1828.

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *Neuroreport*, *16*, 1769–1773.

Arsalidou, M., & Taylor, M. J. (2011). Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage*, *54*, 2382–2393.

Ashcraft, M. H. (1982). The development of mental arithmetic: a chronometric approach. *Developmental Review*, *2*, 213–236.

Ashcraft, M. H. (1992). Cognitive arithmetic—a review of data and theory. *Cognition*, *44*, 75–106.

Ashkenazi, S., Rosenberg-Lee, M., Tenison, C., & Menon, V. (2012). Weak task-related modulation and stimulus representations during arithmetic problem solving in children with developmental dyscalculia. *Developmental Cognitive Neuroscience*, *2*, S152–S166.

Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The codevelopment of skill at and preference for use of retrieval-based processes for solving addition problems: individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology*, *113*, 78–92.

Barrouillet, P., & Lepine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology*, *91*, 183–204.

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 14116–14121.

Barth, H., Starr, A., & Sullivan, J. (2009). Children's mappings of large number words to numerosities. *Cognitive Development*, *24*, 248–264.

Blair, K. P., Rosenberg-Lee, M., Tsang, J. M., Schwartz, D. L., & Menon, V. (2012). Beyond natural numbers: negative number representation in parietal cortex. *Frontiers of Human Neuroscience*, *6*, 7.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*, 746–749.

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Science*, *14*, 277–290.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers counting knowledge. *Developmental Psychology*, *20*, 607–618.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*, 205–228.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*, 425–428.

Butterworth, B. (1999). *The mathematical brain*. London, UK: Macmillan.

Butterworth, B. (2005). Developmental dyscalculia. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 455–457). Hove, UK: Psychology Press.

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Science*, *14*, 534–541.

Cai, W., Chen, T., Ryali, S., Kochalka, J., Li, C. S., & Menon, V. (2016). Causal Interactions within a frontal-cingulate-parietal network during cognitive control: convergent evidence from a multisite-multitask investigation. *Cerebral Cortex*, *26*(5), 2140–2153.

Cai, W., Ryali, S., Chen, T., Li, C. S., & Menon, V. (2014). Dissociable roles of right inferior frontal cortex and anterior insula in inhibitory control: evidence from intrinsic and task-related functional parcellation, connectivity, and response profile analyses across multiple datasets. *Journal of Neuroscience*, *34*, 14652–14667.

Campbell, J. I. (2008). Subtraction by addition. *Memory and Cognition*, *36*, 1094–1102.

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*, 401–406.

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-year-old children. *PLoS Biology*, *4*, e125.

Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009a). The neural development of an abstract concept of number. *Journal of Cognitive Neuroscience*, *21*, 2217–2229.

Cantlon, J. F., Pinel, P., Dehaene, S., & Pelphrey, K. A. (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex*, *21*, 191–199.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009b.) Beyond the number domain. *Trends in Cognitive Science*, *13*, 83–91.

Carey, S. (2004). *On the Origin of Concepts*. New York, NY: Oxford University Press.

Carey, S. (2009). *The Origin of Concepts*. Oxford, UK: Oxford University Press.

Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., & Zilles, K. (2006). The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. *Neuroimage*, *33*, 430–448.

Chang, T. T., Metcalfe, A. W. S., Padmanabhan, A., Chen, T. W., & Menon, V. (2016). Heterogeneous and nonlinear development of human posterior parietal cortex function. *Neuroimage*, *126*, 184–195.

Chang, T. T., Rosenberg-Lee, M., Metcalfe, A. W., Chen, T., & Menon, V. (2015). Development of common neural representations for distinct numerical problems. *Neuropsychologia*, *75*, 481–95.

Cho, S., Metcalfe, A. W., Young, C. B., Ryali, S., Geary, D. C., & Menon, V. (2012). Hippocampal-prefrontal engagement and dynamic causal interactions in the maturation of children's fact retrieval. *Journal of Cognitive Neuroscience*, *24*, 1849–66.

Cho, S., Ryali, S., Geary, D. C., & Menon, V. (2011). How does a child solve 7 + 8? Decoding brain activity patterns associated with counting and retrieval strategies. *Developmental Science*, *14*, 989–1001.

Church, R. M., & Meck, W. H. (1984). The numerical attribute of stimuli. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal Cognition* (pp. 445–464). Hillsdale, NJ: Erlbaum.

Cipolotti, L., Butterworth, B., & Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain*, *114*(Pt 6), 2619–2637.

Cipolotti, L., & Van Harskamp, N. (2001). Disturbances of number processing and calculation In R. S. Berndt (Ed.), *Handbook of Neuropsychology* (pp. 305–331). Amsterdam, the Netherlands: Elsevier Science.

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408–411.

Cohen, L., & Dehaene, S. (1995). Number processing in pure alexia: the effect of hemispheric asymmetries and task demands. *Neurocase*, *1*, 121–137.

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. *Neuron*, *53*, 307–14.

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, *16*, 1348–1355.

Cowan, R., Donlan, C., Shepherd, D. L., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation proficiency and mathematics achievement in elementary school children. *Journal of Educational Psychology*, *103*, 786–803.

Cutini, S., Scatturin, P., Basso Moro, S., & Zorzi, M. (2014). Are the neural correlates of subitizing and estimation dissociable? An fNIRS investigation. *Neuroimage*, *85*(Pt 1), 391–399.

Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Animal Cognition*, *11*, 683–689.

Dantzig, T. (1930). *Number: the Language of Science: A Critical Survey Written for the Cultured Non-mathematician*, New York, NY: Macmillan.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*, 693–700.

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 2157–2162.

De Smedt, B., Holloway, I. D., & Ansari, D. (2011). Effects of problem size and arithmetic operation on brain activation during calculation in children with varying levels of arithmetical fluency. *Neuroimage*, *57*, 771–81.

De Smedt, B., Taylor, J., Archibald, L., & Ansari, D. (2010). How is phonological processing related to individual differences in children's arithmetic skills? *Developmental Science*, *13*, 508–520.

De Smedt, B., Verschaffel, L., & Ghesquiere, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 469–479.

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*, New York, NY: Oxford University Press.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: a logarithmic mental number line. *Trends in Cognitive Science*, *7*, 145–147.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*, 83–120.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*, *33*, 219–250.

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology*, *14*, 218–224.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.

Delazer, M., & Benke, T. (1997). Arithmetic facts without meaning. *Cortex*, *33*, 697–710.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Science*, *11*, 379–86.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, *22*, 1078–1085.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152.

Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning & Behavior*, *25*, 234–244.

Evans, T. M., Kochalka, J., Ngoon, T. J., Wu, S. S., Qin, S. Z., Battista, C., & Menon, V. (2015). Brain structural integrity and intrinsic functional connectivity forecast 6 year longitudinal growth in children's numerical abilities. *Journal of Neuroscience*, *35*, 11743–11750.

Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., . . . Schlaggar, B. L. (2008). The maturing architecture of the brain's default network. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 4028–4032.

Fechner, G. T. (1860). *Elementeder Psychophysik*. Leipzig, Germany: Breitkopf & Hartel.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, *44*, 33–66.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Fias, W., Menon, V., & Szucs, D. (2013). Multiple components of developmental dyscalculia. *Trends in Educational Neuroscience*, *2*, 43–47.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.

Geary, D. C. (1994). *Children's Mathematical Development: Research and Practical Applications*. Washington, DC: American Psychological Association.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Developmental Psychology*, *47*, 1539–1552.

Geary, D. C., Bow-Thomas, C. C., & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: a comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, *54*, 372–391.

Geary, D. C., & Brown, S. C. (1991). Cognitive addition: strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology*, *27*, 398–406.

Geary, D. C., & Hoard, M. K. (2003). Learning disabilities in basic mathematics: deficits in memory and cognition. In J. M. Royer (Ed.), *Mathematical Cognition* (pp. 93–115). Greenwich, CT: Information Age.

Geary, D. C., Hoard, M., Byrd-Craven, J., & Desoto, M. (2004). Strategy choices in simple and complex addition: contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88*, 121–151.

Geary, D. C., Hoard, M., Nugent, L., Byrd-Craven, J., Berch, D., & Mazzocco, M. M. (2007). Strategy use, long-term memory, and working memory capacity. In D. Berch & M. M. Mazzocco (Eds.), *Why Is Math So Hard for Some Children?: The Nature and Origins of Mathematical Learning Difficulties* (pp. 83–105). Baltimore, MD: Paul H. Brookes.

Geary, D. C., & Menon, V. (in press). Fact retrieval deficits in mathematical learning disability: potential contributions of prefrontal-hippocampal functional organization. In M. Vasserman & W. S. MacAllister (Eds.), *The Neuropsychology of Learning Disorders: A Handbook for the Multi-disciplinary Team*. New York, NY: Springer.

Geary, D. C., & Widaman, K. F. (1992). Numerical cognition—on the convergence of componential and psychometric models. *Intelligence*, *16*, 47–80.

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: a critical review. *Acta Psychologica*, *171*, 17–35.

Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.

Gerstmann, J. (1940). Syndrome of finger agnosia, disorientation for right and left, agraphia and acalculia—local diagnostic value. *Archives of Neurology and Psychiatry*, *44*, 398–408.

Giaquinto, M. (1995). Concepts and calculation. *Mathematical Cognition*, *1*, 61–81.

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., . . . Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS ONE*, *8*(6): e67374.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*, 394–406.

Gobel, S. M., Watson, S. E., Lervag, A., & Hulme, C. (2014). Children's Arithmetic development it is number knowledge, not the approximate number sense, that counts. *Psychological Science*, *25*, 789–798.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*, 20–25.

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia*, *47*, 604–608.

Grotheer, M., Herrmann, K. H., & Kovacs, G. (2016). Neuroimaging Evidence of a bilateral representation for visually presented numbers. *Journal of Neuroscience*, *36*, 88–97.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": the approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*, 1457–1465.

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668.

Ham, T., Leff, A., De Boissezon, X., Joffe, A., & Sharp, D. J. (2013). Cognitive control and the salience network: an investigation of error processing and effective connectivity. *Journal of Neuroscience*, *33*, 7091–7098.

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, *341*, 1123–1126.

Heine, A., Tamm, S., Wissmann, J., & Jacobs, A. M. (2011). Electrophysiological correlates of non-symbolic numerical magnitude processing in children: joining the dots. *Neuropsychologia*, *49*, 3238–3246.

Henschen, S. E. (1920). *Klinische und Anatomische Beitraege zur Pathologie des Gehirns*. Stockholm, Sweden: Nordiska Bokhandeln.

Hinton, E. C., Dymond, S., Von Hecker, U., & Evans, C. J. (2010). Neural correlates of relational reasoning and the symbolic distance effect: involvement of parietal cortex. *Neuroscience*, *168*, 138–148.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: the numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 17–29.

Holloway, I. D., Battista, C., Vogel, S. E., & Ansari, D. (2013). Semantic and perceptual processing of number symbols: evidence from a cross-linguistic fMRI adaptation study. *Journal of Cognitive Neuroscience*, *25*(3), 388–400.

Houde, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: an fMRI meta-analysis of 52 studies including 842 children. *Developmental Science*, *13*, 876–885.

Hyde, D. C., Boas, D. A., Blair, C., & Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. *Neuroimage*, *53*, 647–652.

Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: an electrophysiological investigation of small and large number representations. *Journal of Cognitive Neuroscience*, *21*, 1039–1053.

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: but only in children. *Psychonomic Bulletin & Review*, *18*, 1222–1229.

Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., & Delazer, M. (2007). Imaging early practice effects in arithmetic. *Neuroimage*, *36*, 993–1003.

Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstatter, F., Benke, T., Felber, S., & Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *Neuroimage*, *30*, 1365–1375.

Iuculano, T., Tang, J., Hall, C. W., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science*, *11*, 669–680.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: a longitudinal investigation. *Journal of Experimental Child Psychology*, *85*, 103–119.

Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology*, *36*, 763–787.

Kawashima, R., Taira, M., Okita, K., Inoue, K., Tajima, N., Yoshida, H., . . . Fukuda, H. (2004). A functional MRI study of simple arithmetic—a comparison between children and adults. *Cognitive Brain Research*, *18*, 227–233.

Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. *Journal of Neuroscience*, *34*, 9857–9866.

Koehler, O. (1951). The ability of birds to "count." *Bulletin of Animal Behavior*, *9*, 41–45.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends in Cognitive Science*, *17*, 401–412.

Kucian, K., Loenneker, T., Martin, E., & Von Aster, M. (2011). Non-symbolic numerical distance effect in children with and without developmental dyscalculia: a parametric fMRI study. *Developmental Neuropsychology*, *36*, 741–762.

Kucian, K., Von Aster, M., Loenneker, T., Dietrich, T., & Martin, E. (2008). Development of neural networks for exact and approximate calculation: An fMRI study. *Developmental Neuropsychology*, *33*, 447–473.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*, 395–438.

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2016). From "sense of number" to "sense of magnitude"—the role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, 1–62.

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: dissociable systems. *Neuropsychologia*, *41*, 1942–1958.

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development*, *76*, 978–988.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Science*, *17*, 391–400.

Lyons, I. M., & Ansari, D. (2015). Foundations of children's numerical and mathematical skills: the roles of symbolic and nonsymbolic representations of numerical magnitude. *Advances in Child Development and Behavior*, *48*, 93–116.

Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS One*, *6*, e23749.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*, 293–299.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, *1*, 109–121.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology—Animal Behavior Processes*, *9*, 320–334.

Menon, V. (2014). Arithmetic in child and adult brain. In R. Cohen Kadosh & A. Dowker (Eds.), *Handbook of Mathematical Cognition* (pp. 502–530). Oxford, UK: Oxford University Press.

Menon, V., Rivera, S. M., White, C. D., Eliez, S., Glover, G. H., & Reiss, A. L. (2000). Functional optimization of arithmetic processing in perfect performers. *Cognitive Brain Research*, *9*, 343–345.

Metcalfe, A. W., Ashkenazi, S., Rosenberg-Lee, M., & Menon, V. (2013). Fractionating the neural correlates of individual working memory components underlying arithmetic problem-solving skills in children. *Developmental and Cognitive Neuroscience*, *6*, 162–175.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*, 490–502.

Naccache, L., & Dehaene, S. (2001). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11*, 966–974.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*, 1708–1711.

Oberauer, K., Lange, E., & Engle, R. W. (2004). Working memory capacity and resistance to interference. *Journal of Memory and Language*, *51*, 80–96.

Park, J., Hebrank, A., Polk, T. A., & Park, D. C. (2012). Neural dissociation of number from letter recognition and its relationship to parietal numerical processing. *Journal of Cognitive Neuroscience*, *24*, 39–50.

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics.

Journal of Experimental Child Psychology, *88*, 348–367.

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: working memory, phonological ability and numerical competence. *Cognitive Development*, *22*, 165–184.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*, 33–41.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*, 547–555.

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes? *Neuroimage*, *15*, 435–446.

Piffer, L., Agrillo, C., & Hyde, D. C. (2012). Small and large number discrimination in guppies. *Animal Cognition*, *15*, 215–221.

Pinel, P., Dehaene, S., Riviere, D., & Lebihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage*, *14*, 1013–1026.

Prado, J., Mutreja, R., Zhang, H., Mehta, R., Desroches, A. S., Minas, J. E., & Booth, J. R. (2011). Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. *Human Brain Mapping*, *32*, 1932–1947.

Preston, J. L., Molfese, P. J., Frost, S. J., Mencl, W. E., Fulbright, R. K., Hoeft, F., . . . Pugh, K. R. (2016). Print-speech convergence predicts future reading outcomes in early readers. *Psychological Science*, *27*, 75–84.

Price, G. R., Holloway, I., Rasanen, P., Vesterinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, *17*, R1042–3.

Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal-neocortical functional reorganization underlies children's cognitive development. *Nature Neuroscience*, *17*, 1263–1269.

Raizada, R. D. S., Tsao, F. M., Liu, H. M., Holloway, I. D., Ansari, D., & Kuhl, P. K. (2010). Linking brain-wide multivoxel activation patterns to behaviour: examples from language and math. *Neuroimage*, *51*, 462–471.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *Quarterly Journal of Experimental Psychology*, *55*, 1127–1139.

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: the development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*, 480–489.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, *15*, 1779–1790.

Rosenberg-Lee, M., Ashkenazi, S., Chen, T., Young, C. B., Geary, D. C., & Menon, V. (2014). Brain hyper-connectivity and operation-specific deficits during arithmetic problem solving in children with developmental dyscalculia. *Developmental Science*, *18*, 351–372.

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: a coordinate-based meta-analysis. *Neuroimage*, *60*, 830–46.

Rousselle, L., & Noel, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, *102*, 361–395.

Rousselle, L., Palmers, E., & Noel, M. P. (2004). Magnitude comparison in preschoolers: what counts? Influence of perceptual variables. *Journal of Experimental Child Psychology*, *87*, 57–84.

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., . . . Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*, 1250–1255.

Schaffelhofer, S., & Scherberger, H. (2016). Object vision to hand action in macaque parietal, premotor, and motor cortices. *Elife*, 5.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., . . . Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, *27*, 2349–2356.

Shallice, T., & Evans, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex*, *14*, 294–303.

Shum, J., Hermes, D., Foster, B. L., Dastjerdi, M., Rangarajan, V., Winawer, J., Parvizi, J. (2013). A brain area for visual numerals. *Journal of Neuroscience*, *33*, 6709–6715.

Shumaker, R. W., Palkovich, A. M., Beck, B. B., Guagnano, G. A., & Morowitz, H. (2001). Spontaneous use of magnitude discrimination and ordination by the orangutan (*Pongo pygmaeus*). *Journal of Comparative Psychology*, *115*, 385–391.

Siegler, R. S. (1987). The perils of averaging data over strategies: an example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250.

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. New York, NY: Oxford University Press.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237–243.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing Cognitive Competence: New Approaches to Process Modeling* (pp. 31–76). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: how do children know what to do. In C. Sophian (Ed.) *The Origins of Cognitive Skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, *33*, 475–87.

Soltesz, F., Szucs, D., & Szucs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: a developmental study. *Behavioral and Brain Functions*, *6*(13), 13.

Squire, L. R. (1992). Memory and the hippocampus—a synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. In E. Kandel, Y. Dudai, & M. Mayford (Eds.), *Learning and Memory* (pp. 1–14). Cold Spring Harbor, NY: Perspective in Biology Cold Spring Harbor Laboratory Press.

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, *27*, 279–306.

Sridharan, D., Levitin, D. J., & Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 12569–12574.

Starkey, P., & Cooper, R. G. Jr. (1980). Perception of numbers by human infants. *Science*, *210*, 1033–1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*, 97–127.

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 18116–18120.

Supekar, K., & Menon, V. (2012). Developmental maturation of dynamic causal control signals in higher-order cognition: a neurocognitive network model. *PLoS Computational Biology*, *8*(2), e1002374.

Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosenberg-Lee, M., Fuchs, L., & Menon, V. (2013). Neural predictors of individual differences in response to math tutoring in primary-grade school children. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 8230–8235.

Supekar, K., Uddin, L. Q., Prater, K., Amin, H., Greicius, M. D., & Menon, V. (2010). Development of functional and structural connectivity within the default mode network in young children. *Neuroimage*, *52*, 290–301.

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*, 2674–2688.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*, 751–754.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., . . . Morris, R. G. (2007). Schemas and memory consolidation. *Science*, *316*, 76–82.

Tulving, E. (1983). *Elements of Episodic Memory*, New York, NY: Oxford University Press.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132.

Van Loosbroek, E., & Smitsman, A. (1990). Visual perception of numerosity in infancy. *Developmental Psychology*, *26*, 916–922.

Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*, 419–425.

Vogel, S. E., Grabner, R. H., Schneider, M., Siegler, R. S., & Ansari, D. (2013). Overlapping and distinct brain regions involved in estimating the spatial position of numerical and non-numerical magnitudes: an fMRI study. *Neuropsychologia*, *51*, 979–989.

Wimmer, H., Ludersdorfer, P., Richlan, F., & Kronbichler, M. (2016). Visual Experience shapes orthographic representations in the visual word form area. *Psychological Science*, *27*, 1240–1248.

Wu, S. S., Meyer, M. L., Maeda, U., Salimpoor, V., Tomiyama, S., Geary, D. C., & Menon, V. (2008). Standardized assessment of strategy use and working memory in early mental arithmetic performance. *Developmental Neuropsychology*, *33*, 365–393.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749–50.

Wynn, K. (1998). Psychological foundations of number: numerical competence in human infants. *Trends in Cognitive Science*, *2*, 296–303.

Zamarian, L., & Delazer, M. (2014). Arithmetic learning in adults—evidence from brain imaging. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition*. Oxford, UK: Oxford University Press.

CHAPTER 7

Development of Scientific Thinking

CORINNE ZIMMERMAN AND DAVID KLAHR

WHY STUDY SCIENTIFIC THINKING?

For almost a century, psychologists interested in cognitive development have devised empirical investigations to uncover the origins and trajectory of scientific thinking and have explored a variety of methods for enriching children's understanding of scientific procedures and concepts. The study of scientific thinking is particularly appealing to cognitive and developmental psychologists not only because of science's cultural value but also because of the inherent importance and challenge of investigating scientific thinking and the paradox of "the child as scientist." In addition, science educators are interested in the topic because of its obvious relevance for improving science instruction.

Cultural Value

Science and technology have had profound effects on human culture. Scientific thinking has enhanced the ability of human beings to understand, predict, and control the natural forces that shape our world. The term "scientific literacy" refers to the skills required by citizens in a scientifically advanced society (Bybee, 2015; DeBoer, 2000). Students, citizens, and policy makers need to understand how to investigate, evaluate, and comprehend science content (e.g., climate change, evolution, astronomy, disease), processes (e.g., how to test hypotheses effectively), and products (e.g., from evaluating data about the most effective cancer treatments to the possibility of space colonization). In addition to the "factual" aspects of scientific knowledge and scientific procedures, the well-being of a society depends on a widespread appreciation of the value of science (e.g., the necessity of evidence-based decisions about policies, practices, and programs). A report from the National Research Council (NRC; 2010) argued that science is *the* discipline that should be used to convey the skills required for the 21st-century workforce, such as nonroutine problem-solving, adaptability, complex communication skills, self-management, and systems thinking.

Inherent Importance and Difficulty

Although scientific thinking has its roots in "everyday thinking," it is much more complex, highly structured, and refined. There has been a rich mythology about the importance, on one hand, and the intractability, on the other, of studying the psychological processes that lead to scientific discovery. With respect to intractability, Einstein once mused, "I am not sure whether there can be a way of really understanding the miracle of thinking" (cited in Wertheimer, 1959, p. 227).

If Einstein's bewilderment were correct, then this chapter could conclude right here. However, in a more optimistic and constructive reflection, Einstein also said:

> The whole of science is nothing more than a refinement of everyday thinking. It is for this reason that the critical thinking of the physicist cannot possibly be restricted to the examination of concepts of his own specific field. He cannot proceed without considering critically a much more difficult problem, the problem of analyzing the nature of everyday thinking. (1936, p. 59)

In the decades since Einstein had this remarkable insight, cognitive scientists have made substantial advances in understanding the "nature of everyday thinking." At the same time, there has been a rich and active line of research on the cognitive and social processes involved in scientific inquiry and discovery. Thus, the very difficulty of studying this complex, multilayered, and socially impactful topic provides a strong justification for the endeavor.

The Paradox of the "Child as Scientist"

Both children and scientists are described as "naturally curious" with an inherent and enthusiastic approach to finding out about the world. Research shows that infants and young children have some of the precursors and abilities needed to engage in formal scientific thinking (Chaille & Britain, 1991; Gopnik, Meltzoff, & Kuhl, 2000). Yet older children and adults struggle with scientific thinking tasks; they can be unsystematic, ignore and misinterpret evidence, try to prove what they already believe to be true, and design uninformative experiments (to name just a few of the documented difficulties). Although U- and inverted-U-shaped curves are common in developmental psychology, there is more to this story than early and late competence. By systematically studying age-appropriate tasks across the life span, we have learned a lot about the extent to which the child-as-scientist view is supported and the extent to which, from preschool to college, people have deep and well-entrenched misconceptions about both the content and processes of science.

Educational Relevance

The study of basic cognitive processes involved in scientific thinking has obvious and important implications for science education across the curriculum, from prekindergarten through college. For example, Piaget's theory of cognitive development was very influential in the design and development of science curricula from the late 1950's through the end of the 20th century (Blake & Pope, 2008; Elkind, 1972; Kamii & DeVries, 1993; Klahr, 2012; Metz, 1995, 1997). And even as the field of cognitive development has distanced itself from much of Piaget's theoretical edifice, many current educational practitioners, from kindergarten through college level, continue to base much of their instruction on Piagetian stages of cognitive development. Science education (in the United States in particular) has undergone remarkable change within the past several decades (DeBoer, 2000). Recently, in the United States, national organizations such as the National Research Council have been focused on evidence-based efforts to improve science education. These efforts have included an explicit acknowledgment of the basic psychological research on scientific thinking (e.g., NRC, 2000, 2007, 2008). The "Taking Science to School K–8" report (NRC, 2007) emphasized and summarized several important emergent themes in recent research in the learning sciences and cognitive development. For example, with respect to summarizing what research has told us about how science is learned, the "surprisingly sophisticated" nature of children's thinking is acknowledged, as is the ability of children to engage in a variety of reasoning processes that represent the precursors to mature scientific thinking. The role of adults, teachers, and learning experiences is emphasized as a way to turn children's "rich but naïve understandings of the natural world" (p. 3) into proficient skills needed for engaging in investigation, evidence evaluation, explanation, argumentation, and discourse as science students and as scientifically literate adults. Recommendations about best practices for supporting this transition have been proposed (NRC, 2007, 2008) that acknowledge children's existing and developing capabilities, focus on core knowledge areas and skills, and promote the opportunity for the students to engage in the practices of science.

The plan for this chapter is as follows. We (1) define scientific thinking by drawing on two relatively distinct lines of inquiry; (2) present a taxonomy to categorize these lines of inquiry and the cognitive skills involved; (3) describe illustrative examples of experimental studies of scientific thinking; and (4) summarize what has been learned about the similarities and differences between children's scientific thinking and mature scientific thinking.

WHAT IS SCIENTIFIC THINKING?

Science, as a human endeavor, can be approached as an individual, social, and cultural activity. At the individual level, scientific thinking shares many characteristics with other forms of problem solving and reasoning (Klahr, Matlen, & Jirout, 2013; Zimmerman & Croker, 2013). It can be described further as a specific type of intentional information or knowledge seeking (Kuhn, 2011). Curiosity emerges early and spontaneously in children (Jirout & Klahr, 2012). However, before this innate curiosity can address scientific issues effectively, it must be refined and shaped by instruction through deliberate activities, such as exploring, asking questions, testing hypotheses, engaging in inquiry, and evaluating evidence (Jirout & Zimmerman, 2015; Morris, Croker, Masnick, & Zimmerman, 2012). An additional defining feature of mature scientific thinking involves metacognitive and metastrategic knowledge—the ability to reflect on the process of knowledge acquisition and the changes that result from engaging in scientific activities (Kuhn, 2011). Metacognitive skills are evident in children's emerging theory of mind skills. Children must learn where beliefs about the world around them come from, that others may have different beliefs, that beliefs can be more or less certain, and, in particular, that beliefs may be formed on the basis of inference or from evidence (e.g., Sodian & Wimmer, 1987). This broad and encompassing definition provides a context in which to summarize a very wide range of investigations, from studies of young children making observations in a school classroom (e.g., Chinn & Malhotra, 2002) to descriptions of a research team in a laboratory discussing the results of a set of experiments (e.g., Dunbar, 1995).

Psychological studies of scientific thinking have taken several forms, including historical accounts and case studies of individual scientists or groups of scientists (see Klahr, 1994; Klahr & Simon, 1999 for review) and computational models of the cognitive processes underlying scientific problem solving and discovery. (See Shrager & Langley, 1990, for overview.) In this chapter, we focus on psychological studies of participants in simulated discovery contexts. Participants from some characteristic population (e.g., school children, college students, scientists) are presented with problem-solving situations that isolate one or more essential aspects of "real-world" science. The "thing to be discovered" can range from something as simple and arbitrary as a "rule" that the experimenter has in mind (e.g., Wason, 1960), to something as complex as the physics of an artificial universe (Mynatt, Doherty, & Tweney, 1977) or the mechanisms of genetic inhibition (Dunbar, 1993). The advantage of this approach is that it enables the researcher to exert some experimental control over participants' prior knowledge and complete control over the "state of nature." Most important, this approach enables the researcher to observe the dynamic course of scientific discovery processes in great detail.

Conceptual Thinking in Science

The beginning of psychological research on scientific thinking has been widely attributed to Jean Piaget's meeting with Albert Einstein in 1928. Einstein was curious about the developmental origins of fundamental physical concepts. Piaget began a line of research that included children's understanding of scientific concepts and their skills for investigating the world and credited Einstein for inspiration (Piaget 1946). Piaget investigated children's developing thinking processes about time, speed, distance, number, movement, velocity, living things, people, space, mathematics, logic, morality, physical causality, and psychology. The legacy of his steadfast focus on children's early understanding of causes and effects in the natural world cannot be overstated. Piaget's influence was eloquently summarized by John Flavell (1996) and continues to be reiterated (Klahr, 2012).

Following on in this tradition of studying scientific thinking, researchers have examined the concepts that children and adults hold in the various domains of science, such as biology (e.g., Carey, 1985; Hatano & Inagaki, 2013), chemistry (Calik & Ayas, 2005; Garnett, Garnett, & Hackling, 1995), and physics (e.g., Chi, Feltovich, & Glaser, 1981). Numerous examples of specific concepts within these domains have been studied. For example, conceptual development has been studied in astronomy (Vosniadou & Brewer, 1994); buoyancy, current, and bubbles (Tenenbaum, Rappolt-Schlichtmann, & Zanger, 2004); earth and space concepts (Sackes, 2015); ecology (Zimmerman & Cuddington, 2007); evolution (Emmons & Kelemen, 2015; Samarapungavan & Wiers, 1997); genetics (Echevarria, 2003); gravity (Hood, 1998); and life science concepts (Akerson, Wieland, & Fouad, 2015).

The focus in this line of research has been to identify and describe the mental models or domain-specific theories that children and adults hold about scientific phenomena and the progression of changes that these models undergo with experience or instruction. Here, scientific thinking is studied by asking participants to use their conceptual knowledge about particular scientific phenomena to answer questions and reason about novel scenarios. Two classic studies illustrate this type of scientific thinking. To probe children's understanding of the shape of the earth, Vosniadou and Brewer (1992) asked 6- to 11-year-olds factual questions, such as "What is the shape of the earth?," and questions that would differentiate children's conceptualizations, such as "What is above/below the earth?" and "Can you fall off the edge [of the earth]?" These responses, along with drawings, uncovered a variety of alternative mental models that varied in internal consistency (e.g., disc earth, hollow sphere, rectangular earth). In the domain of genetics, Clough and Driver (1986) asked early adolescents to reason about situations of genetic inheritance involving the offspring of humans or animals with acquired traits (e.g., a gardener with rough skin, a mouse that had lost its tail). Developmental trends in beliefs about inheritance and accompanying explanations were evident from 12 to 16 years of age. These studies are designed to assess children's existing understanding of scientific concepts. Other studies, such as work by Kelemen, Emmons, Seston, and Ganea (2014), have demonstrated that children's deep scientific misconceptions (e.g., about natural selection) can be remediated by relatively brief engagement in well-designed explicit instruction. An interesting aspect of this line of research is that in many cases (such as the Vosniadou & Brewer, 1992, work mentioned earlier), it appears that "ontogeny recapitulates phylogeny." That is, young children's initial "flat-earth" beliefs are similar to those of prescientific cultures, and only over time (and with instruction) do children realize that the earth is a solid sphere.

The vast number of possible concepts that can be investigated makes it difficult to summarize the findings from this literature adequately, because of the domain-specific nature of such research. For example, in 2009, Reinders Duit, a researcher at University of Kiel, compiled a bibliography of research studies on conceptual change in science with more than 8,400 entries. (For the most recent update on this bibliography, see http://archiv.ipn.uni-kiel.de/stcse/.) Consider, for example, a single chapter on children's understanding of physical science concepts. Hadzigeorgiou (2015) reviewed studies about children's ideas about matter, heat, temperature, evaporation, condensation, the water cycle, forces, motion, floating, sinking, electricity, and light. Each of these topics can be further unpacked to constituent subcomponents (e.g., electricity concepts include current, voltage, charge, electrons, resistance, and circuits, to name a few).

In the most current science standards in the United States, the NRC (2012) identified three dimensions that represent "a broad set of expectations for students in science" (p. 1), including disciplinary core ideas, crosscutting concepts, and scientific and engineering practices. The core ideas are restricted to the traditional sciences (i.e., physical, life, earth/space). The crosscutting concepts, in contrast, are broadly applicable to several domains of science and include ideas such as causality, patterns, time, feedback, analogy, and equilibrium. There is a rich literature on what individuals understand about the concept of causality across the life span, from infancy studies (e.g., Baillargeon, 2004) through childhood (e.g., Gopnik, Sobel, Schulz & Glymour, 2001; Piaget, 1974) and adulthood (for a review, see Koslowski & Masnick, 2010). Although less work has been done on the other domain-general crosscutting concepts, there are a few exceptions. Swanson (2015) examined students' understanding of patterns (as a process or behavior) that can underlie physical, social, or psychological phenomena (e.g., threshold, equilibration, and oscillation) in a middle-school science course called "The Patterns Class."

Procedural Thinking in Science

In addition to studies addressing what children know about science, there is a substantial literature on how children acquire that knowledge. Here, too, Piaget led the way in his decision to study not only what children know about the world at various stages of their development but also, and perhaps more important, the methods and processes that they use to acquire, integrate, and refine this knowledge. This second line of research examines the development of scientific thinking by focusing on the scientific practices of observing, asking questions, conducting experiments, evaluating evidence, constructing models, and generating explanations. In this line of investigation, "scientific thinking is something people *do*, not something they *have*" (Kuhn, 2011, p. 498, emphasis in original).

The distinction between the products and processes of scientific thinking is reflected in the often-distinct research programs that developmental psychologists have undertaken. It is possible to develop a line of research about children's knowledge about various content (e.g., astronomy, biology) without needing to be concerned about their investigation skills (i.e., how they came to know it). For example, Vosniadou and Brewer's (1992) study of children's conceptions about the shape of the earth is a canonical study of scientific knowledge, but it does not address children's knowledge-producing activities. Studies of such knowledge-producing investigation skills have utilized two main strategies. One is to reduce the reliance on (or interference by) conceptual knowledge by creating tasks that reduce the role of prior knowledge (e.g., Siegler & Liebert, 1975; Wason, 1960). The other is to examine these knowledge types in a more integrated way, motivated by the fact that concepts and procedures are intertwined in authentic scientific thinking.

TAXONOMY FOR CATEGORIZING EXPERIMENTAL STUDIES OF SCIENTIFIC THINKING

Our taxonomy utilizes the two principle features that have been used previously to characterize scientific thinking: domain-specific concepts and domain-general procedures (Klahr, 1994; Klahr, Zimmerman, & Jirout, 2011; Zimmerman, 2000, 2007). In addition, Table 7.1 extends this taxonomy to incorporate some of the scientific and engineering practices that have been identified by the NRC's (2012) national science education standards. The cognitive processes identified by Klahr and Dunbar (1988) are included in the first column. The scientific practices that map on to these cognitive processes are included in the second column. In the third and fourth columns, we distinguish between domain-general and domain-specific knowledge because some studies focus on specific content whereas others use simple or abstract contexts that do not require much (or any!) domain-specific knowledge. For example, a typical domain-specific study might investigate what children understand about the domain of chemistry by asking them to reason about processes such as dissolution and chemical change (Calik & Ayas, 2005). Domain-general studies, in contrast, focus on broadly applicable reasoning processes that can be investigated in arbitrary and abstract forms, such as Bruner's classic concept learning tasks (Bruner, Goodnow, & Austin, 1956) or Wason's famous 2–4–6 task (Wason, 1960).

Some preliminaries are in order. Some studies might fit neatly into a single cell, but for the majority of this literature (described in more detail in the next section), even studies that focus on a particular cognitive process or scientific practice have to traverse more than one cell. For example, in order to examine experimentation skills, participants must set up (or select from a predefined set of choices) an experiment to address a particular hypothesis. Likewise, the evaluation of evidence must be done in some context and requires a consideration of either the experiment that produced it or the hypothesis that it is meant to support or refute.

EMPIRICAL INVESTIGATIONS OF THE DEVELOPMENT OF SCIENTIFIC THINKING

Klahr and Dunbar's (1988) Scientific Discovery as Dual Search (SDDS) model will serve as the general framework for organizing

Table 7.1 Taxonomy for Categorizing Experimental Studies of Scientific Thinking, with Representative Examples of Each Type

		Type of Knowledge	
Cognitive Processes	**Science Practices**	**Domain Specific**	**Domain General**
Forming and Refining Hypotheses (Hypothesis Space Search)	Asking questions	**A1** Metz (2004) Samarapungavan et al. (2008)	**D1** Chouinard (2007) Jirout & Klahr (2017)
	Developing and using models	**A2** Vosniadou & Brewer (1992, 1994) Lehrer & Schauble (2004)	**D2** Raghavan & Glaser (1995) Greca & Moreira (2000) Wu & Puntambekar (2012)
Investigation Skills (Experiment Space Search)	Planning and carrying out investigations	**B** Tschirgi (1980) Schwichow et al. (2016b)	**E** Siegler & Liebert (1975) Kuhn & Ho (1980)
Evaluating Evidence	Analyzing and interpreting data/evidence	**C1** Penner & Klahr (1996b) Masnick et al. (2016)	**F1** Shaklee & Paszek (1985) Masnick & Morris (2008)
	Constructing explanations	**C2** Gelman & Kremer (1991) Carey & Spelke (1994) Hatano & Inagaki (2013)	**F2** Mynatt et al. (1978) Schunn & Klahr (1993) Kelemen et al. (2014)

NOTE: The cognitive processes categories and knowledge types are adapted from Klahr and Dunbar (1988), Klahr (1994), and Klahr and Carver (1995). The science practices are from the *Framework for K–12 Science Education* (NRC, 2012). Two additional NRC (2012) practices—"Engaging in argument from evidence" and "Obtaining, evaluating, and communicating information"—are beyond the scope of this chapter. Similarly, "Mathematical and computational thinking" is a separate research literature despite its obvious connection to authentic scientific practice (but see Chapter 6 of this volume, "Development of Mathematical Thinking"). In each cell, we note one or more typical exemplars, although any given study might traverse several cells (e.g., Penner & Klahr, 1996a); additional work representative of each cell is described in the main text.

illustrative empirical findings to be discussed. (See Table 7.1.) The SDDS framework captures the complexity and the cyclical nature of the process of scientific discovery and includes both inquiry skills and conceptual change. (See Klahr, 2000, for a detailed discussion.) The top-level categories of the model include the three major cognitive components of scientific discovery: searching for hypotheses, searching for experiments (or investigations more generally), and evidence evaluation. The studies to be reviewed involve one or more of these three processes. SDDS is an extension of a classic model of problem solving from the field of cognitive science (Newell & Simon, 1972; Simon & Lea, 1974) and explains how people carry out problem solving in varied science contexts, from simulated inquiry to professional scientific practice. The fundamental aspects of SDDS are (a) the concept of two distinct but closely related "problem spaces": a space of hypotheses and a space of experiments; and (b) coordinated search in these two problem spaces.

Individuals begin inquiry tasks with some existing or intuitive ideas, or perhaps no ideas at all about how particular variables influence an outcome. Given some set of possible variables (i.e., independent variables) and asked to determine their effect on an outcome (i.e., the dependent variable), participants negotiate the process by coordinating search in the set of possible hypotheses and the set of possible experiments. Experiments are conducted to determine the truth status of the current hypothesis or to decide among a set of competing hypotheses. Experiments also may be conducted to generate enough data to be able to propose a hypothesis (as might be the case when one has little or no prior knowledge). Evidence then is evaluated so that inferences can be made whether a hypothesis is correct or incorrect (or, in some cases, that the evidence is inconclusive). Depending on the complexity of the task, the number of variables, and the amount of time on task, these processes may be repeated several times as an individual negotiates search in the hypothesis and experiment spaces and makes inferences based on the evaluation of self-generated evidence. Factors such as task domain, amount of prior knowledge, and the perceived goal of the task influence how these cognitive processes are deployed.

Searching the Hypothesis Space

Of the three cognitive processes of SDDS, search in the hypothesis space has the most in common with conceptual thinking in science, as it typically involves a search of relevant domain-specific knowledge as represented in the hypothesis space. When one is engaged in inquiry or investigation activities, hypothesis-space search is instantiated in the service of the scientific practices of asking questions and developing or using models (NRC, 2012).

Asking Questions and Curiosity

Asking questions is one of the foundational process skills of scientific practice (NRC, 2012). Older students often believe that the goal of science is to demonstrate what is already known (Kuhn, 2005) or to see if something "works" or to invent things (Carey, Evans, Honda, Jay, & Unger, 1989). However, asking questions for which the answer is not yet known is a crucial element of inquiry that students must learn (Kuhn & Dean, 2005). They must learn not only how to ask "good" questions but also that question-asking is a defining feature of science. An essential precursor to asking good questions is curiosity (Jirout & Klahr, 2012; Klahr, Zimmerman, & Jirout, 2011). The fundamental importance of curiosity in science education is indicated by its nearly universal inclusion across a variety of highly influential science curricula, educational standards, and assessment goals (American Association for the Advancement of Science, 1993; National Education Goals Panel, 1993, 1995; NAEYC, 2012; NRC, 2000). Curiosity is the desire or motivation to explore and ask questions. Specifically, we define curiosity as the preferred level of uncertainty—or the amount of uncertainty that will lead to deliberate question-asking or exploratory behavior (Jirout & Klahr, 2012, 2017; Jirout & Zimmerman, 2015).

Simple problem-solving tasks that require question-asking have been used for investigating children's ability to recognize specific instances of uncertainty and to evaluate information. Referential tasks assess children's general ability to ask categorical questions ("Is it an animal?," "Does it bark?") that will help them to identify a target from a group of possibilities (e.g., one picture from an array of pictures). When children are given the opportunity to ask a question to figure out which object is hidden in a box before guessing, they are correct on about five of the six trials; if they are told to guess what is in the box without being allowed to ask a question, their accuracy is at chance (Chouinard, 2007). Thus, children can determine which questions to ask to address uncertainty, but they also can use information that the answers to their questions yield to resolve it.

Research on conceptual thinking in science has used various methodologies (drawings, standardized interviews, reasoning scenarios) to uncover children's understanding of various phenomena. At this intersection with process skills, there are examples of children's question-asking in particular domains. For example, Greif, Kemler Nelson, Keil, and Gutierrez (2006) investigated young children's ability to ask domain-specific questions on a structured task. Children were instructed to ask questions about unfamiliar objects and animals, which they were able to do—averaging 26 questions asked across 12 pictures. Many questions were quite general, such as "What is it?" Other questions, however, showed that children recognized and understood that different questions should be asked of the different categories (i.e., objects and animals). Children tended to ask more function questions about the objects (i.e., what can I do with this?), whereas the unfamiliar animals prompted questions about category membership, food choices, and locations. Samarapungavan, Mantzicopoulos, and Patrick (2008) explored whether kindergarteners could generate meaningful questions about biological topics, such as structure, function, and growth with respect to living things in general and about monarch butterflies in particular. Given a supportive inquiry unit, the majority of students were found to be proficient at asking questions that were meaningful and biologically relevant. Similarly, Metz (2004) examined the specific questions that children in the second and fourth grades asked about cricket behavior a supportive inquiry-based curriculum. With appropriate scaffolding, all pairs of students were able to formulate appropriate and relevant research questions. Moreover, these children were able to generate fairly sophisticated and domain-specific categories of researchable questions (e.g., comparison of cricket behavior under different conditions).

Developing and Using Models

Developing and using models constitutes a "signature practice of the sciences" (Quellmalz, Timms, Silberglitt, & Buckley, 2012, p. 366) that is becoming increasingly emphasized in science education and science assessment (Clement, 2000; Lehrer & Schauble, 2000, 2012; NRC, 2012). The use of models to support theory building, argumentation, and explanation is common in science and engineering (Nersessian, 2008) and has been documented in sociohistorical analyses of practicing scientists (e.g., Thagard, 2000; Tweney, 2002). Scientists use physical models (e.g., Watson and Crick's model of the structure of DNA), drawings, and schematic representations (e.g., Faraday's sketches of electromagnetic toruses or Darwin's tree of life). In both science and science education, the ability to develop and use models is becoming increasingly easier (and thus, by extension, easier to be more sophisticated) due to the scaffolding provided by computers and computer simulations. Simulation models can be used to learn about and investigate phenomena that are "too large, too small, too fast, too slow, or too dangerous to study in classrooms" (Quellmalz et al., 2012, p. 367).

Schunn and Klahr's (1995, 1996) investigation of participants' explorations of a complex computer microworld (see Klahr, 2000, chapter 7, for additional details) led them to propose that as tasks become more rich, complex, and authentic, the SDDS model of scientific thinking should include the search of two additional problem spaces. As well as hypotheses and experiments, scientific discovery requires the search of abstract data representations (or "models") and a space of experimental paradigms at a level of abstraction above the specific instantiation of an experiment. The data representation (or model) space is used to select attributes of the data under consideration.

This process may involve attention to regularities of the data set or may result from analogy to existing knowledge (e.g., the superficial similarity between an atom and a solar system). Experimental paradigms function as general-purpose templates that can be deployed in multiple specific, but isomorphic, contexts. One example of a simple experimental paradigm that is located in the paradigm space is the concept of ensuring that a variable hypothesized to be critical is manipulated. The SDDS model, and the extension involving a total of four search spaces, represents a domain-general model of the scientific thinking processes involved in scientific discovery.

In science education, there are numerous domain-general and domain-specific examples of the instantiation of such model-based practices. A concept like variability is one that can be applied in several different domains (Lehrer & Schuable, 2004). Domain-specific examples are varied and include learning about decomposition in a first-grade classroom (Ero-Tolliver, Lucas, & Schauble, 2013), evolution in elementary school (Keleman et al., 2014; Lehrer & Schauble, 2012), ecosystems in sixth grade (Lehrer, Schauble, & Lucas, 2008), and biomechanics of the human elbow in college (Penner, Lehrer, & Schauble, 1998).

Searching the Experiment Space

Given a hypothesis (or a set of competing hypotheses), one usually needs to design an experiment, and the design of that experiment can be construed as a problem to be solved, via search in a problem space (Newell & Simon, 1972). Of course, experimentation is just one of several types of legitimate scientific inquiry processes (see Lehrer, Schauble, & Petrosino, 2001), but here we focus on the substantial body of literature on the development of experimentation skills. A "solution" to problem-solving search in the experiment space is an experiment that assesses the truth or falsity of the current hypothesis. The privileged status of experimentation in science education is indicated by its inclusion in science standards (NRC, 2012; NGSS Lead States, 2013). In the next sections, first we describe research on the developmental precursors of experimentation skills, then we present studies in which participants are engaged in the full cycle of experimentation.

Early Experimentation Skills

Science education for young children tends to focus on investigation skills such as observing, describing, comparing, and exploring (NAEYC, 2012; National Science Teachers Association, 2007). Even though, as noted earlier, few contemporary researchers in cognitive development accept the assertions of Piaget's stage theory (e.g., Inhelder & Piaget 1958; Piaget, 1970), it often has been used to justify waiting until adolescence before attempting to teach science process skills (French & Woodring, 2013; Metz, 1995). However, an accumulation of evidence about human learning (e.g., NRC, 2000) has resulted in a more nuanced story about the development of experimentation and investigation skills and the extent to which well-designed instruction can accelerate that development (NRC, 2007). The identification of causal factors in the world via experimentation involves the coordination of several component processes: identifying and manipulating variables and observing and measuring outcomes. Not until the later school years, after extended instruction, scaffolding, and practice, can children successfully coordinate all of these steps (e.g., Kuhn, Black, Keselman, & Kaplan, 2000). Several studies have examined the precursors of the later ensemble of experimentation skills, however, and we turn to them next.

In her classic study, Tschirgi (1980) presented children and adults with a variety of everyday problem-solving situations (e.g., baking cakes, making paper airplanes) that involved a positive or negative outcome and several potential causal variables, such as "John baked a cake using honey, white flour, and butter, and it turned out terrible" or "Susan made a paper airplane and it turned out great." The character would propose a hypothesis about a variable that may have caused the outcome (e.g., "John thinks that the honey made it taste bad" in the cake story). The participant in the study then would be asked to select one of three options in order to help the character (John) test the hypothesis. In the vary-one-thing-at-a-time (VOTAT) option, the proposed variable was changed, but the others were kept the same (e.g., bake another cake, with everything the same except the sweetener: use sugar instead of honey). This strategy would produce an unconfounded experiment. In the hold-one-thing-at-a-time (HOTAT) option, the hypothesized variable was kept the same, but the other variables were changed (e.g., bake another cake with the same sweetener but change the type of flour and shortening). The change-all option consisted of changing all of the variables (bake a cake with sugar, whole-wheat flour, and margarine). All participants were more likely to select the HOTAT strategy when the outcome was positive. That is, the presumed causal variable was held constant (consistent with a confounded experiment) to maintain the positive outcome. For a negative outcome, the logically correct VOTAT strategy (consistent with a controlled experiment) was chosen more frequently than HOTAT or change-all, suggesting that participants were searching for the one variable to change in order to eliminate the negative outcome. Although second and fourth graders were more likely to select the change-all strategy for the negative outcomes (likely as a way to eliminate all possible offending variables), all participants were influenced by the desire to reproduce good effects and eliminate bad effects by choosing a strategy based on pragmatic outcomes (rather than logical grounds).

Croker and Buchanan (2011) used a task similar to Tschirgi's but included contexts for which 3.5- to 11-year-olds held strong prior beliefs (e.g., the effect of cola versus milk on dental health). For all age groups, there was an interaction of prior belief and outcome type. The logically correct VOTAT strategy was more likely to be selected under two conditions: (a) when the outcome was positive (i.e., healthy teeth) and consistent with prior belief, or (b) when the outcome was negative (i.e., unhealthy teeth) and inconsistent with prior belief. Even the youngest children were influenced by the context and the plausibility of the domain-specific content of the situations that they were reasoning about.

Sodian, Zaitchik, and Carey (1991) investigated the extent to which children in the early school years understand that one can use an experiment to engage in hypothesis testing. Their classic mouse house task presents children with a situation that pits "finding out" against "producing an effect." First- and second-grade children were presented with a challenge in which they had to figure out whether their home contained a large mouse or a small mouse. Children were shown "mouse houses" in which they could put some food that mice like. One house had a door through which either a large or a small mouse could pass. The other house had a door that only a small mouse could traverse. In the "find out" condition, the children were asked to decide which house should be used to determine the size of the mouse (i.e., to test a hypothesis). Of course, if they use the house with the small door and the food is gone in the morning, they know that they have a small mouse. If the food remains, they have a large (and now hungry!) mouse. Importantly, Sodian et al. had a second condition, the "feed" condition, in which children were asked what house to use if they wanted to make sure that the mouse would get fed, no matter what its size. If a child can distinguish between the goals of testing a hypothesis with an experiment versus generating an effect (i.e., feeding the mouse), then he or she should select the small house in the find-out condition and the large house in the feed condition. Sodian et al. found that children as young as 6 could distinguish between a conclusive and inconclusive experimental test of a simple hypothesis when provided with the two mutually exclusive and exhaustive hypotheses or experiments.

Chen and Klahr (2008) used tasks isomorphic to those used by Sodian et al. (1991) in a training study aimed at determining the extent to which kindergartners through second graders could distinguish conclusive from inconclusive experimental tests. In addition, they used different types of feedback (implicit, verbal, physical demonstration) to show that children could transfer strategies for selecting/generating an experimental test of a hypothesis over delays up to 24 months. Piekny and Maehler (2013) used the mouse house task with preschoolers (4- and 5-year-olds) and school children (7-, 9-, and 11-year-olds). It was not until age 9 that children scored significantly above chance and not until age 7 (a year later than in the Sodian et al. study) that children showed a recognition of, and justification for, conclusive or inconclusive tests of a hypothesis.

Klahr, Fay, and Dunbar (1993) investigated experimentation skills by presenting subjects (third- and sixth-grade children and adults) with a programmable toy robot, in which participants first mastered most of the basic commands. They were then challenged to find out how a "mystery key" worked by writing and then running programs that included the mystery key. In order to constrain the "hypothesis space," participants were provided with various hypotheses about the mystery key that were correct or incorrect. Some examples of what the mystery key might do include (a) repeat the whole program *n* times, (b) repeat the last step *n* times, or (c) repeat the last *n* steps once. Some of these hypotheses were deemed highly plausible (i.e., likely to be correct), and others were implausible. When a presented hypothesis was plausible, all participants set up experiments to demonstrate the correctness of the hypothesis. When given an implausible hypothesis to test, adults and some sixth graders proposed a plausible rival hypothesis and set up an experiment that would discriminate between the two. The third graders also proposed a plausible rival hypothesis but got sidetracked in the attempt to demonstrate that the rival plausible hypothesis was correct. Klahr et al. identified two useful heuristics that participants used: (a) design experiments that produce informative and interpretable results, and (b) attend to one feature at a time. The third- and sixth-grade children were much less likely than the adult participants to restrict the search of possible experiments to those that were informative.

Bullock and Ziegler (1999) collected longitudinal data on participants, starting when they were age 8 and following them through to age 12. They examined the process skills required for experimentation, using separate assessments to tease apart an understanding of experimentation from the ability to produce controlled experiments. When the children were 8 years old, they were able to recognize a controlled experimental test. The ability to produce a controlled experiment at levels comparable to adults did not occur until the children were 12 years old. This study provides additional support for the idea that young children are able to understand the "logic" of experiments long before they are able to produce them.

When task demands are reduced—such as simple story problems or when one can select (rather than produce) an experimental test—even young children show competence with rudimentary science process skills. Children, like adults, are sensitive to the context and the content of what is being reasoned about.

Planning and Carrying Out Investigations

For older children and adults, much of the research on the development of investigation skills involves presenting participants with a multivariable causal system, such as physical apparatus or a computer simulation. The participants' goal is to investigate the system so as to identify the causal and noncausal variables in the system; they propose hypotheses, make predictions, plan and conduct experiments, collect and evaluate evidence, make inferences, and draw conclusions in the form of either new or updated knowledge. For example, Schauble's (1996) participants conducted experiments in hydrodynamics, where the goal was to determine which variables had an effect on boat speed. Participants could vary the depth of the canal and the size, shape, and weight of the boat.

A foundational science process skill is the control-of-variables strategy (CVS), which is a domain-general skill (Chen & Klahr, 1999). The fundamental goal of an experiment is to unambiguously identify causal factors and their effects, and the essential procedure for doing this is to contrast conditions that differ only with the respect to the variable whose causal status is under investigation. Procedurally, CVS includes the ability to create experiments in which conditions differ with respect to only a single contrasting variable as well as the ability to recognize confounded and unconfounded experiments. Conceptually, CVS involves the ability to make appropriate inferences from the results of unconfounded experiments (e.g., that only inferences about the causal status of the variable being tested are warranted) as well as an awareness of "the inherent indeterminacy of confounded experiments" (Chen & Klahr, 1999, p. 1098). The conceptual aspects of CVS are relevant for argumentation and reasoning about causality in science and everyday life, as CVS includes an understanding of the invalidity of evidence from confounded experiments (or observations) and the importance of comparing controlled conditions (Kuhn, 2005). Thus, CVS is relevant to broader educational and societal goals, such as inquiry, reasoning skills, and critical thinking.

Mastery of CVS is required for successful inquiry learning as it enables students to conduct their own informative investigations. However, without instruction, students and even adults have poor inquiry skills (e.g., Kuhn, 2007; for review, see Zimmerman & Croker, 2013). Siler and Klahr (2012) identified the various "misconceptions" that students have about controlling variables. Typical mistakes include (a) designing experiments that vary the wrong (or "nontarget") variable, (b) varying more than one variable, or (c) not varying anything between the contrasted experimental conditions (i.e., overextending the "fairness" idea so both conditions are identical).

A recent meta-analysis of CVS interventions (Schwichow, Croker, Zimmerman, Höffler, & Härtig, 2016a) summarized the results of 72 studies. Possible moderators of the overall effect size included design features (e.g., quasi-experimental versus experimental studies), instructional features (e.g., use of demonstrations), training features (e.g., use of hands-on experiences), and assessment features (e.g., test format). Of the various instructional features coded for, only two were found to be effective: (a) interventions that induced a cognitive conflict and (b) teacher demonstrations of good experimental design. In this context, a teacher draws attention to a particular (confounded) comparison and asks what conclusions can be drawn about the effect of a particular variable. For example, to return to the cake baking example described earlier, a teacher might note that although the cake made with butter, whole-wheat flour, and sugar tasted much better than the cake made with margarine, white flour, and sugar, one could not tell for sure if the effect was due to the type of flour or the type of sweetener. Because the comparison was confounded, with two possible causal factors, either one of these potential causes might have determined the outcome.

Cognitive conflict is induced in students by drawing attention to a current experimental procedure or interpretation of data; the teacher attempts to get the student to notice that the comparison is confounded or that the conclusion is invalid or indeterminate (Adey & Shayer, 1990). Interestingly, the cognitive conflict technique often is presented via a demonstration by the teacher, so additional research is necessary to disentangle the unique effects of these two instructional techniques (Schwichow et al., 2016a). Other instructional techniques that often are presumed to be important, such as the need for "hands-on" engagement with experimental materials, did not have an impact on student learning of CVS. In a follow-up to the meta-analysis, Schwichow, Zimmerman, Croker, and Härtig (2016b) determined that it is important for there to be a match between the way students learn CVS and the test format used to assess the extent to which they have learned it.

Evaluating Evidence

The goal of most experiments is to produce evidence that bears on a hypothesis, and once that evidence is generated, it must be interpreted. (We say "most" here because, in some cases, scientists may perform experiments in the absence of any clearly articulated hypothesis, just to get a "feel" for the nature of the phenomenon.) The final cognitive process and scientific practices we discuss are those that enable people to evaluate and explain how evidence relates to the hypothesis that inspired it. Evidence evaluation is the part of the cycle of inquiry aimed at determining whether the result of an experiment (or set of experiments) is sufficient to reject or accept a hypothesis under consideration (or whether the evidence is inconclusive) and to construct possible explanations for how the hypothesis and evidence are related.

Interpreting evidence always is done in the context of prior belief. However, to minimize the effects of (widely varying) prior knowledge, early investigations of the evidence evaluation process used tasks that deemphasized it. Wason's (1960) famous 2–4–6 task was conceived as a "knowledge lean" task to simulate the process of evaluating evidence to test and revise hypotheses. In brief, the participant is given an exemplar (2–4–6) of a general rule for numerical triads and asked to hypothesize the rule that governs the sequence of digits. The cycle includes hypothesis formation, generation of a new triad to test the hypothesis, evidence (feedback that the participant-generated triad is, or is not, an exemplar of the rule), and hypothesis revision; this cycle continues until enough evidence has accumulated to discover the rule. Traditionally, the experimenter chooses a very general rule, such as "3 integers in ascending order," but most participants begin the task by (mistakenly) hypothesizing a much narrower rule (e.g., "even numbers increasing by 2"). Research using this task became a cottage industry in cognitive psychology—in fact, Wason (1960) has been cited more than 1000 times—with many interpretations and reinterpretations of performance (e.g., Klayman & Ha, 1987).

Interestingly, there are few developmental studies using the 2–4–6 task: "[C]hildren do not know enough about numerical relationships to make the mistakes, so typical in the task with adults, and hence they often hit on the rule immediately. They have not yet learnt to erect their own obstacles against finding it." (Wason & Johnson-Laird, 1972, p. 217)

As will be illustrated, much of the literature on evidence evaluation demonstrates that both children and adults have a tendency to "erect their own obstacles" to knowledge acquisition.

Evaluating Patterns of Evidence

One method of examining the developmental precursors of skilled evidence evaluation with children involves presenting them with pictorial representations of potential causes and effects. These are often simple representations like those between types of food and health (e.g., Kuhn, Amsel, & O'Loughlin, 1988) or plant treatment (e.g., sun, water) and plant health (Amsel & Brock, 1996). The pictures may represent perfect covariation between cause and effect, partial covariation, or no covariation. This cognitive skill is facilitated by the metacognitive ability to make a distinction between a hypothesis and the evidence to support a hypothesis (Kuhn, 2005, 2011).

In their classic study, Ruffman, Perner, Olson, and Doherty (1993) presented 4- to 7-year-old children with simple story problems involving one potential cause (e.g., type of food: red or green) and an outcome (tooth loss). A "faked evidence task" was used to determine whether children could form different hypotheses based on varying patterns of evidence. For example, children would be shown that green food perfectly covaries with tooth loss: This situation represents the "real evidence." Next, the evidence was tampered with; anyone who was unaware of the original pattern would be led to believe that red food causes tooth loss (i.e., the "faked evidence"). Children were asked to interpret which hypothesis the faked evidence supported. The key advantage of this type of task is that it is diagnostic with respect to whether a child can make a distinction between a hypothesis and a pattern of evidence to support a hypothesis. This task requires children to understand that their own hypothesis would be different from that of a story character who saw only the faked evidence. When considering the responses to both the initial hypothesis-evidence task and the faked-evidence task, only the 5- to 7-year-olds performed above chance level. Partial covariation evidence was used to determine if 5- to 7-year-olds could form hypotheses based on patterns of evidence. When considering both hypothesis-evidence and faked-evidence questions, only the performance of the 6- and 7-year olds was above chance level. Most children understood that veridical versus faked evidence would lead to different beliefs and that a newly formed hypothesis could be used to generalize to future cases.

Ruffman et al. (1993) showed that some of the very basic prerequisite evidence-evaluation skills required for scientific thinking are present as early as 6 years of age. In follow-up research, Koerber, Sodian, Thoermer, and Nett (2005) examined the performance of 4- to 6-year-olds on a variety of evidence evaluation tasks to examine whether existing causal beliefs influence evidence evaluation in the preschool years. In situations where there are no strong prior beliefs and the outcomes are equally plausible, preschoolers correctly interpreted perfect and partial covariation evidence. Preschoolers had difficulty, however, with evidence that contradicted prior plausible beliefs; this finding is consistent with the performance of both older children and adults on scientific thinking tasks (Zimmerman & Croker, 2013). Although young children demonstrate some of the precursors to more advanced evidence-evaluation skills, they too are susceptible to the influences of prior beliefs and considerations of the plausibility of what is being evaluated.

Beginning with the foundational work of Kuhn et al. (1988), we know that the process of revising and acquiring knowledge on the basis of evidence is highly influenced by the prior knowledge that a participant brings to the task. Evaluating evidence is guided by an assessment of the plausibility of a hypothesized cause; we make judgments about the world in ways that "make sense" or are consistent with what we already know about how things work. Plausibility is a known constraint in belief formation and revision (Holland, Holyoak, Nisbett, & Thagard, 1986) and is a domain-general heuristic that is used to guide the choice of which hypotheses to test and which experiments to run (Klahr et al., 1993). Because the strength of existing beliefs and assessments of plausibility are both considered when evaluating evidence, children and adults often choose to maintain their prior beliefs rather than changing them to be in line with newly acquired evidence (e.g., Chinn & Brewer, 1998; Chinn & Malhotra, 2002). A common finding is that it is generally more difficult to integrate evidence that disconfirms a prior causal belief (which involves restructuring one's belief system) than it is to integrate evidence that disconfirms a prior noncausal belief (which involves incorporating a newly discovered causal relation). For example, children and adults have robust physics misconceptions about weight, mass, and density, and these misconceptions influence the evaluation of evidence in tasks that involve the motion (e.g., falling, sinking, rolling) of objects. In the case of sinking objects, it is difficult to give up the belief that weight matters, but it is easy to add the belief that shape (sphere versus cube) speeds up or slows down an object based on firsthand evidence (Penner & Klahr, 1996b). Other research shows pervasive difficulties with revising knowledge on the basis of evidence, even when that evidence is generated and observed directly (rather than being provided by researchers; e.g., Chinn & Malhotra, 2002; Renken & Nunez, 2010).

Analyzing and Interpreting Data

The NRC (2012) science standards include the scientific practice of analyzing and interpreting data. The standards note that "scientific investigations produce data that must be analyzed in order to derive meaning [. . .] data do not speak for themselves" (p. 51). An inescapable aspect of empirical research is that all measurements in the physical world include some degree of error, and children must learn how to deal with it. Masnick, Klahr, and Morris (2007) describe the challenge for the young scientist:

> A young child eagerly awaits the day when she will pass the 100 cm minimum height requirement for riding on the "thriller" roller coaster at her local amusement park. She regularly measures her height on the large-scale ruler tacked to her closet door. As summer approaches, she asks her parents to measure her every week. A few weeks ago she measured 98 cm, last week 99.5 cm, but today only 99.0 cm. Disappointed and confused, when she gets to school she asks the school nurse to measure her, and is delighted to discover that her height is 100.1 cm. Success at last! But as she anticipates the upcoming annual class excursion to the amusement park, she begins to wonder: what is her real height? And more importantly, what will the measurement at the entrance to the roller coaster reveal? Why are all the measurements different, rather than the same? Because she is a really thoughtful child, she begins to speculate about whether the differences are in the thing being measured

> (i.e., maybe her height really doesn't increase monotonically from day to day) or the way it was measured (different people may use different techniques and measurement instruments when determining her height). (p. 3)

Although the processes associated with understanding and interpreting error and data variability draw heavily on mathematical reasoning, and therefore are beyond the scope of this chapter, a few studies capture the intersection of analyzing quantitative data and identifying sources of error. Masnick and Morris (2008) examined how the characteristics of measurement data, such as sample size and variability within the data set (e.g., magnitude of differences, presence of outliers) influenced conclusions drawn by third and sixth graders and adults. Participants were shown data sets with plausible cover stories (e.g., testing new sports equipment) and were asked to indicate what conclusions could be drawn and their reasons. Third and sixth graders had rudimentary skills in detecting trends, overlapping data points, and the magnitude of differences. Sixth graders had developing ideas about the importance of variability and the presence of outliers for drawing conclusions from data. At all ages, participants were more confident of conclusions based on larger samples of observations.

Masnick, Klahr, and Knowles (2017) explored how adults and children (aged 9–11) responded to (a) variability in the data collected from a series of simple experiments and (b) the extent to which the data were consistent with their prior hypotheses. Participants conducted experiments in which they generated, recorded, and interpreted data to identify factors that affect the period of a pendulum. In Study 1, several children and most adults used observed evidence to revise their initial understanding, but participants were more likely to change incorrect noncausal beliefs to causal beliefs than the reverse. In Study 2, participants were oriented toward either an "engineering" goal (to produce an effect) or a "science" goal (to discover the causal structure of the domain) and were presented with variable data about potentially causal factors. Science goals produced more belief revision than engineering goals. Numerical data, when presented in context and with appropriate structure, can help children and adults reexamine their beliefs and initiate and support the process of conceptual change and robust scientific thinking.

Constructing Explanations

The NRC's (2012) Framework for Science Education emphasizes the importance of scientific theories and explanations: "The goal for students is to construct logically coherent explanations of phenomena that incorporate their current understanding of science, or a model that represents it, and are consistent with the available evidence" (p. 52). Scientific explanations typically are constructed after investigations that produce evidence that is to be evaluated and ultimately explained.

Much has been written in the scientific thinking literature about the ability to differentiate between evidence (i.e., data, observation, patterns) and the explanation or theory that purports to account for that evidence. In particular, Kuhn's (1989, 2005, 2011) research has emphasized that mature scientific thinking requires the cognitive and metacognitive skills to differentiate between evidence and the theory or explanation for that evidence. Kuhn argued that effective coordination of evidence and theory depends on three metacognitive abilities: (a) The ability to encode and represent evidence and theory separately, so that relations between them can be recognized; (b) the ability to treat theories or explanations as independent objects of thought (i.e., rather than a representation of "the way things are"); and (c) the ability to recognize that theories can be false and explanations flawed, and, having recognized that possibility, to assess the evidence in order to determine whether the theory is true or false. These metacognitive abilities are necessary precursors to sophisticated scientific thinking and represent one of the ways in which children, adults, and professional scientists differ.

As noted previously, children are inclined to notice and respond to causal events in the environment; even infants and young children have been shown to have rudimentary understanding of cause and effect (Bullock & Gelman, 1979; Piaget, 1929). Keil's (2006; Keil & Wilson, 2000) work on the nature of explanation in general indicates that children and adults alike have a propensity to generate explanations. We often privilege causal explanations, which are arguably quite important in scientific thinking. Koslowski's (1996, 2012, 2013) research showed that people are good at noticing evidence for the covariation between events in the world, but there is a tendency to make only causal inferences when the link can be explained with a causal mechanism. Participants consider or generate plausible causal mechanisms to explain the relationship between potential causes and their effects. Similarly, if a plausible causal mechanism exists to explain why a cause and effect should be linked, it is difficult to let go of that belief. Therefore, we see across many types of scientific thinking tasks that both children and adults have a strong tendency to maintain beliefs rather than change them based on evidence (e.g., Chinn & Brewer, 1998; Chinn & Malhotra, 2002) because the strength of existing beliefs, assessments of plausibility, casual mechanisms, and alternative causal mechanisms are all potentially salient and brought to bear when reasoning (Koslowski, Marasia, Chelenza, & Dublin, 2008).

CONCLUSION: REVISITING THE CHILD AS SCIENTIST

In this review, we have illustrated the two main approaches to studying the development of scientific thinking. One line of research has focused on the content of science: what children and adults think about various science concepts in the traditionally defined disciplines of science. The second has focused on the processes or procedures of science: how children and adults ask questions, solve problems, conduct investigations, and evaluate evidence to revise their explanations about how the world works.

The natural and social worlds comprise the laboratory of both the scientist and the child. Scientific thinking is considered "a hallmark intellectual achievement of the human species" (Feist, 2006, p. ix). As developmental psychologists, however, we are interested in the factors that influence the origins and growth of scientific thinking, from the child in a science classroom through to the practicing scientist. In scientific thinking, two developmental endpoints get emphasized—the child and the scientist. Children have been likened to scientists; scientists are said to have the curiosity of young children. But as developmental research from the past century has shown, there is a lot that goes on in between. As is the case with other academic skills, such as reading and mathematical thinking, scientific thinking is highly mediated both culturally and educationally. Researchers have come to acknowledge that the cognitive processes and the set of scientific practices that must be coordinated in mature scientific thinking require practice and are developed within a social context and with the aid of cultural tools (Lemke, 2001; Zimmerman & Croker, 2014). Importantly, the ability to reflect metacognitively on the process of knowledge acquisition and change is a hallmark of fully developed scientific thinking (Kuhn, 2005).

Science educators now recognize that students should be exposed to learning experiences that reflect how real science is conducted and communicated, and educational reform has been aimed at how best to engage students in developmentally appropriate authentic inquiry and argumentation (NRC, 2007, 2008, 2011, 2012). Unlike other basic cognitive skills (e.g., attention, perception, memory), scientific thinking does not "routinely develop" (Kuhn & Franklin, 2006, p. 974); that is, scientific thinking does not emerge independent of culture and cultural tools. Metacognitive abilities are necessary precursors to sophisticated scientific thinking but also represent one of the ways in which children, adults, and professional scientists differ. In order for children to go beyond demonstrating the correctness of their existing beliefs (e.g., Dunbar & Klahr, 1989), they must develop and practice meta-level competencies. With metacognitive control over the processes involved, children can change what they believe based on evidence. In doing so, they are not only aware that they are changing a belief; they also know why they are changing a belief. Thus, sophisticated scientific thinking involves the cognitive processes involved in asking questions, forming and refining hypotheses, conducting investigations, developing models, designing experiments, evaluating evidence, and constructing explanations as well as a meta-level awareness of when, how, and why one should engage in these practices.

REFERENCES

Adey, P., & Shayer, M. (1990). Accelerating the development of formal thinking in middle and high school students. *Journal of Research in Science Teaching*, *27*, 267–285.

Akerson, V. L., Weiland, I. S., Fouad, K. (2015). Children's ideas about life science. In K. C. Trundle & M. Sackes (Eds.), *Research in Early Childhood Science Education* (pp. 99–123). Dordrecht, the Netherlands: Springer.

American Association for the Advancement of Science. (1993). *Benchmarks for Scientific Literacy: A Project 2061 Report*. New York, NY: Oxford University Press.

Amsel, E., & Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, *11*, 523–550.

Baillargeon, R. (2004). Infants' reasoning about hidden objects: evidence for event-general and event-specific expectations. *Developmental Science*, *7*, 391–424.

Blake, B., & Pope, T. (2008). Developmental psychology: incorporating Piaget's and Vygotsky's theories in classrooms. *Journal of Cross-Disciplinary Perspectives in Education*, *1*, 59–67.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A Study of Thinking*. London, UK: Chapman & Hall.

Bullock, M., & Gelman, R. (1979). Preschool children's assumptions about cause and effect: temporal ordering. *Child Development*, *50*(1), 89–96.

Bullock, M., & Ziegler, A. (1999). Scientific reasoning: developmental and individual differences. In F. E. Weinert & W. Schneider (Eds.), *Individual Development from 3 to 12: Findings from the Munich Longitudinal Study* (pp. 38–54). Cambridge, UK: Cambridge University Press.

Bybee, R. (2015). Scientific literacy. In R. Gunstone (Ed.), *Encyclopedia of Science Education* (pp. 944–947). Dordrecht, the Netherlands: Springer.

Calik, M., & Ayas, A. (2005). A comparison of level of understanding of eighth-grade students and science student teachers related to selected chemistry concepts. *Journal of Research in Science Teaching*, *42*, 638–667.

Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, MA: MIT Press.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). "An experiment is when you try it and see if it works": A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, *11*, 514–529.

Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 169–200). Cambridge, UK: Cambridge University Press.

Chaille, C., & Britain, L. (1991). *The Young Child as Scientist*. New York, NY: Harper Collins.

Chen, Z., & Klahr, D. (1999). All other things being equal: children's acquisition of the control of variables strategy. *Child Development*, *70*, 1098–1120.

Chen, Z., & Klahr, D. (2008). Remote transfer of scientific-reasoning and problem-solving strategies in children. *Advances in Child Development and Behavior*, *36*, 420.

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Chinn, C. A., & Brewer, W. F. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching*, *35*, 623–654.

Chinn, C. A., & Malhotra, B. A. (2002). Children's responses to anomalous scientific data: how is conceptual change impeded? *Journal of Educational Psychology*, *94*, 327–343.

Chouinard, M. M. (2007). Children's questions: a mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*, 1–129.

Clement, J. (2000). Model-based learning as a key research area for science education. *International Journal of Science Education*, *22*, 1041–1053.

Clough, E. E., & Driver, R. (1986). A study of consistency in the use of students' conceptual frameworks across different task contexts. *Science Education*, *70*, 473–496.

Croker, S., & Buchanan, H. (2011). Scientific reasoning in a real world context: the effect of prior belief and outcome on children's hypothesis testing strategies. *British Journal of Developmental Psychology*, *29*, 409–424.

DeBoer, G. E. (2000). Scientific literacy: another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, *37*, 582–601.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, *17*, 397–434.

Dunbar, K. (1995). How scientists really reason: scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The Nature of Insight* (pp. 365–395). Cambridge, MA: MIT Press.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery processes. In D. Klahr & K. Kotovsky (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon* (pp. 109–143). Hillsdale, NJ: Erlbaum.

Echevarria, M. (2003). Anomalies as a catalyst for middle school students' knowledge construction and scientific reasoning during science inquiry. *Journal of Educational Psychology*, *95*, 357–374.

Einstein, A. (1936). Physics & reality. Reprinted in A. Einstein (1950), *Out of My Later Years*. New York, NY: Philosophical Library,

Elkind, D. (1972). Piaget and science education. *Science and Children*, *10*, 9–12.

Emmons, N. A., & Kelemen, D. (2015). Young children's acceptance of within-species variation: implications for essentialism and teaching evolution. *Journal of Experimental Child Psychology*, *139*, 148–160.

Ero-Tolliver, I., Lucas, D., & Schauble, L. (2013). Young children's thinking about decomposition: early modeling entrees to complex ideas in science. *Research in Science Education*, *43*, 2137–2152.

Feist, G. J. (2006). *The Psychology of Science and the Origins of the Scientific Mind*. New Haven, CT: Yale University Press.

Flavell, J. (1996). Piaget's legacy. *Psychological Science*, *7*, 201–203.

French, L. A., & Woodring, S. D. (2013). Science education in the early years. In O. N. Saracho & B. Spodek (Eds.), *Handbook of*

Research on the Education of Young Children (3rd ed., pp. 179–196). New York, NY: Routledge.

Garnett, P. J., Garnett, P. J., & Hackling, M. W. (1995). Students' alternative conceptions in chemistry: a review of research and implications for teaching and learning. *Studies in Science Education*, *25*, 69–95.

Gelman, S. A., & Kremer, K. E. (1991). Understanding natural cause: children's explanations of how objects and their properties originate. *Child Development*, *62*, 396–414.

Greca, I. M., & Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education*, *22*, 1–11.

Greif, M. L., Kemler Nelson, D. G., Keil, F. C., & Gutierrez, F. (2006). What do children want to know about animals and artifacts? Domain-specific requests for information. *Psychological Science*, *17*, 455.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (2000). *The Scientist in the Crib: What Early Learning Tells Us about the Mind*. New York, NY: Harper Collins.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*, 620–629.

Hadzigeorgiou, Y. (2015). Young children's ideas about physical science concepts. In K. C. Trundle & M. Sackes (Eds.), *Research in Early Childhood Science Education* (pp. 67–97). Dordrecht, the Netherlands: Springer.

Hatano, G., & Inagaki, K. (2013). *Young Children's Thinking about the Biological World*. New York, NY: Psychology Press.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction*. Cambridge, MA: MIT Press.

Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science*, *1*, 59–63.

Inhelder, B., & Piaget, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. New York, NY: Basic Books.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: in search of an operational definition of an elusive concept. *Developmental Review*, *32*, 125–160.

Jirout, J., & Klahr, D. (2017). "Questions—and some answers—about young children's questions." Unpublished manuscript.

Jirout, J., & Zimmerman, C. (2015). Development of science process skills in the early childhood years. In K. C. Trundle & M. Sackes (Eds.), *Research in Early Childhood Science Education* (pp. 143–165). Dordrecht, the Netherlands: Springer.

Kamii, C., & DeVries, R. (1993). *Physical Knowledge in Preschool Education: Implications of Piaget's Theory*. New York, NY: Teachers College Press.

Keil, F. C. (2006). Explanation and Understanding. *Annual Review of Psychology*, *57*, 227–254.

Keil, F. C., & Wilson, R.A. (Eds.) (2000). *Explanation and Cognition*. Cambridge, MA: MIT Press.

Kelemen, D., Emmons, N., Seston, R. & Ganea, P. (2014). Young children can be taught basic natural selection using a picture storybook intervention. *Psychological Science*, *25*, 893–902.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis-testing. *Psychological Review*, *94*, 211–228.

Klahr, D. (1994). Searching for cognition in cognitive models of science. *Psycoloquy*, *5*(94).

Klahr, D. (2000). *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge, MA: MIT Press.

Klahr, D. (2012). Beyond Piaget: a perspective from studies of children's problem solving abilities. In A. Slater & P. Quinn (Eds.), *Refreshing Developmental Psychology: Beyond the Classic Studies* (pp. 56–70). London, UK: Sage.

Klahr, D., & Carver, S. (1995). Scientific thinking about scientific thinking. *Monographs of the Society for Research in Child Development*, *60*(4), 137–151.

Klahr, D., & Dunbar, K. (1988). Dual search space during scientific reasoning. *Cognitive Science*, *12*, 1–48.

Klahr, D., Fay, A., & Dunbar, K. (1993). Heuristics for scientific experimentation: a developmental study. *Cognitive Psychology*, *25*, 111–146.

Klahr, D., Matlen, B. & Jirout, J. (2013). Children as scientific thinkers. In G. Feist & M. Gorman (Eds.), *Handbook of the Psychology of Science* (pp. 223–248). New York, NY: Springer.

Klahr, D., & Simon, H. A. (1999). Studies of scientific discovery: complementary approaches and convergent findings. *Psychological Bulletin*, *125*, 524–543.

Klahr, D., Zimmerman, C., & Jirout, J. (2011). Educational interventions to advance children's scientific thinking. *Science*, *333*, 971–975.

Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology*, *64*, 141–152.

Koslowski, B. (1996). *Theory and Evidence: The Development of Scientific Reasoning*. Cambridge, MA: MIT Press.

Koslowski, B. (2012). How explanation makes information evidentially relevant. In R. Proctor & J. Capaldi (Eds.), *The Psychology of Science: Implicit and Explicit Processes* (pp. 112–136). New York, NY: Oxford University Press.

Koslowski, B. (2013). Scientific reasoning: explanation, confirmation bias, and scientific practice. In G. Feist & M. Gorman (Eds.), *Handbook of the Psychology of Science* (pp. 151–192). New York, NY: Springer.

Koslowski, B., Marasia, J., Chelenza, M., & Dublin, R. (2008). Information becomes evidence when an explanation can incorporate it into a causal framework. *Cognitive Development*, *23*, 472–487.

Koslowski, B., & Masnick, A. M. (2010). Causal reasoning. In U. Goswami (Ed.), *Handbook of Child Cognitive Development* (2nd ed., pp. 377–393). Oxford, UK: Wiley-Blackwell.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*, 674–689.

Kuhn, D. (2005). *Education for Thinking*. Cambridge, MA: Harvard University Press.

Kuhn, D. (2007). Reasoning about multiple variables: control of variables is not the only challenge. *Science Education*, *91*, 710–726.

Kuhn, D. (2011). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Handbook of Childhood Cognitive Development* (2nd ed., pp. 497–523). Oxford, UK: Wiley-Blackwell.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The Development of Scientific Thinking Skills*. Orlando, FL: Academic Press.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*, *18*, 495–523.

Kuhn, D., & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science*, *16*, 866–870.

Kuhn, D., & Ho, V. (1980). Self-directed activity and cognitive development. *Journal of Applied Developmental Psychology*, *1*, 119–130.

Kuhn, D., & Franklin, S. (2006). The second decade: what develops (and how)? In W. Damon & R. Lerner (Series Eds.), D. Kuhn & R. Siegler (Vol. Eds.), *Handbook of Child Psychology: Vol. 2. Cognition, Perception, and Language* (6th ed. pp. 517–550). Hoboken, NJ: Wiley.

Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. *Journal of Applied Developmental Psychology*, *21*, 39–48.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, *41*, 635–679.

Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education*, *96*, 701–724.

Lehrer, R., Schauble, L., & Lucas, D. (2008). Supporting development of the epistemology of inquiry. *Cognitive Development*, *23*, 512–529.

Lehrer, R., Schauble, L., & Petrosino, A. J. (2001). Reconsidering the role of experiment in science education. In K. Crowley, C. Schunn, & T. Okada (Eds.), *Designing for Science: Implications from Everyday, Classroom, and Professional Settings* (pp. 251–277). Mahwah, NJ: Erlbaum.

Lemke, J. L. (2001). Articulating communities: sociocultural perspectives on science education. *Journal of Research in Science Teaching*, *38*, 296–316.

Masnick, A., M., Klahr, D., & Knowles, E. R. (2017). Data-driven belief revision in children and adults. *Journal of Cognition and Development*, *18*(1), 87–109.

Masnick, A. M., Klahr, D., & Morris, B. J. (2007). Separating signal from noise: children's understanding of error and variability in experimental outcomes. In M. Lovett & P. Shaw (Eds.), *Thinking with Data* (pp. 3–26). Mawah, NJ: Erlbaum.

Masnick, A. M., & Morris, B. J. (2008). Investigating the development of data evaluation: the role of data characteristics. *Child Development*, *79*, 1032–1048.

Metz, K. E. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research*, *65*, 93–127.

Metz, K. E. (1997). On the complex relation between cognitive developmental research and children's science curricula. *Review of Educational Research*, *67*, 151–163.

Metz, K. E. (2004). Children's understanding of scientific inquiry: their conceptualization of uncertainty in investigations of their own design. *Cognition & Instruction*, *22*, 219–290.

Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In H. Kloos, B. J. Morris, & J. L. Amaral (Eds.), *Current Topics in Children's Learning and Cognition* (pp. 61–82). Rijeka, Croatia: InTech.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: an experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, *29*, 85–95.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, *30*, 395–406.

NAEYC [National Association for the Education of Young Children]. (2012). All criteria document. Retrieved from http://www.naeyc.org/files/academy/file/AllCriteriaDocument.pdf

National Education Goals Panel. (1993). *Promises to Keep: Creating High Standards for American Students*. Washington, DC: National Education Goals Panel.

National Education Goals Panel. (1995). *The 1995 National Education Goals Report: Building a Nation of Learners*. Washington, DC: National Education Goals Panel.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academies Press.

National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K–8*. Washington, DC: National Academies Press.

National Research Council. (2008). *Ready Set Science: Putting Research to Work in K–8 science Classrooms*. Washington, DC: National Academies Press.

National Research Council. (2010). *Exploring the Intersection of Science Education and 21st Century Skills: a Workshop Summary*. Washington, DC: National Academies Press.

National Research Council. (2012). *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press.

National Science Teachers Association. (2007). *NSTA Position Statement: The Integral Role of Laboratory Investigations in Science Instruction*. Retrieved from http://www.nsta.org/about/positions/laboratory.aspx

Nersessian, N. J. (2008). *Creating Scientific Concepts*. Cambridge, MA: MIT Press.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. New York, NY: Prentice Hall.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, by States*. Washington, DC: National Academies Press.

Penner, D. E., & Klahr, D. (1996a). The interaction of domain-specific knowledge and domain-general discovery strategies: a study with sinking objects. *Child Development*, *67*, 2709–2727.

Penner, D. E., & Klahr, D. (1996b). When to trust the data: further investigations of system error in a scientific reasoning task. *Memory & Cognition*, *24*, 655–668.

Penner, D. E., Lehrer, R., & Schauble, L. (1998). From physical models to biomechanics: a design-based modeling approach. *Journal of the Learning Sciences*, *7*(3–4), 429–449.

Piaget, J. (1929). *The Child's Conception of the World*. London, UK: Kegan Paul,

Piaget, J. (1969). *The Child's Conception of Time*. New York, NY: Ballantine. (Originally published 1946.)

Piaget, J. (1970). *Genetic Epistemology*. New York, NY: Columbia University Press.

Piaget, J. (1974). *Understanding Causality*. New York, NY: Norton. (Originally published 1971.)

Piekny, J., & Maehler, C. (2013). Scientific reasoning in early and middle childhood: the development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology*, *31*, 153–179.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching*, *49*, 363–393.

Raghavan, K., & Glaser, R. (1995). Model-based analysis and reasoning in science: the MARS curriculum. *Science Education*, *79*(1), 37–61.

Renken, M. D., & Nunez, N. (2010). Evidence for improved conclusion accuracy after reading about rather than conducting a belief-inconsistent simple physics experiment. *Applied Cognitive Psychology*, *24*, 792–811.

Ruffman, T., Perner, J., Olson, D. R., & Doherty, M. (1993). Reflecting on scientific thinking: children's understanding of the hypothesis-evidence relation. *Child Development*, *64*, 1617–1636.

Sackes, M. (2015). Young children's ideas about earth and space science concepts. In K. C. Trundle & M. Sackes (Eds.), *Research in Early Childhood Science Education* (pp. 35–65). Dordrecht, the Netherlands: Springer.

Samarapungavan, A., Mantzicopoulos, P., & Patrick, H. (2008). Learning science through inquiry in kindergarten. *Science Education*, *92*, 868–908.

Samarapungavan, A., & Wiers, R. W. (1997). Children's thoughts on the origin of species: a study of explanatory coherence. *Cognitive Science*, *21*, 147–177.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32*, 102–119.

Schunn, C. D. & Klahr, D. (1993). Self vs. other-generated hypotheses in scientific discovery. *Proceedings of Fifteenth Annual Conference of the Cognitive Science Society* (pp. 900–905). Hillsdale, NJ: Erlbaum.

Schunn, C. D., & Klahr, D. (1995). A 4-space model of scientific discovery. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 106–111). Mahwah, NJ: Erlbaum.

Schunn, C. D., & Klahr, D. (1996). Integrated yet different: logical, empirical, and implementational arguments for a 4-space model of inductive problem solving. In G. Cottrell (Ed.), *Proceedings of the Eighteenth Annual Meeting of the Cognitive Science Society* (pp. 25–26). Mahwah, NJ: Erlbaum.

Schwichow, M., Croker, S., Zimmerman, C., Höffler, T., & Härtig, H. (2016a). The control-of-variables strategy: a meta-analysis. *Developmental Review*, *39*, 37–63.

Schwichow, M., Zimmerman, C., Croker, S., & Härtig, H. (2016b). What students learn from hands-on activities. *Journal of Research in Science Teaching*, *53*, 980–1002.

Shaklee, H., & Paszek, D. (1985). Covariation judgment: systematic rule use in middle childhood. *Child Development*, *56*, 1229–1240.

Shrager, J., & Langley, P. (1990). *Computational Models of Scientific Discovery and Theory Formation*. San Mateo, CA: Morgan Kaufman.

Siegler, R. S., & Liebert, R. M. (1975). Acquisition of formal scientific reasoning by 10- and 13-year-olds: designing a factorial experiment. *Developmental Psychology*, *11*, 401–402.

Siler, S., & Klahr, D. (2012). Detecting, classifying and remediating children's explicit and implicit misconceptions about experimental design. In R. W. Proctor & E. J. Capaldi (Eds.), *Psychology of Science: Implicit and Explicit Processes* (pp. 137–180). New York, NY: Oxford University Press.

Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: a unified view. In L. W. Gregg (Ed.), *Knowledge and Cognition* (pp. 105–128). Hillsdale, NJ: Erlbaum.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, *58*, 424–433.

Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, *62*, 753–766.

Swanson, H. L. (2015). "Leveraging students' prior knowledge in attaining deep structural understanding of domain general models" (Doctoral dissertation). http://eprints.cdlib.org/uc/item/4z0914hz

Tenenbaum, H. R., Rappolt-Schlichtmann, G., & Zanger, V. (2004). Children's learning about water in a museum and in the classroom. *Early Childhood Research Quarterly*, *19*, 40–58.

Thagard, P. (2000). *How Scientists Explain Disease*. Princeton, NJ: Princeton University Press.

Tschirgi, J. E. (1980). Sensible reasoning: a hypothesis about hypotheses. *Child Development*, *51*, 1–10.

Tweney, R. D. (2002). Epistemic artifacts: Michael Faraday's search for the optical effects of gold. In L. Magnani & N. J. Nersessian (Eds.), *Model-Based Reasoning: Science, Technology, Values* (pp. 287–303). New York, NY: Kluwer Academic/Plenum Press.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: a study of conceptual change in childhood. *Cognitive Psychology*, *24*, 535–585.

Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science*, *18*, 123–183.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content*. London: B. T. Batsford.

Wertheimer, M. (1959). *Productive Thinking*. New York, NY: Harper & Row. (Originally published 1945.)

Wu, H. K., & Puntambekar, S. (2012). Pedagogical affordances of multiple external representations in scientific processes. *Journal of Science Education and Technology*, *21*, 754–767.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, *20*, 99–149.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*, 172–223.

Zimmerman, C., & Croker, S. (2013). Learning science through inquiry. In G. Feist & M. Gorman (Eds.), *Handbook of the Psychology of Science* (pp. 49–70). New York, NY: Springer.

Zimmerman, C., & Croker, S. (2014). A prospective cognition analysis of scientific thinking and the implications for teaching and learning science. *Journal of Cognitive Education and Psychology*, *13*, 245–257.

Zimmerman, C., & Cuddington, K. (2007). Ambiguous, circular and polysemous: students' definitions of the "balance of nature" metaphor. *Public Understanding of Science*, *16*, 393–406.

CHAPTER 8

Theory of Mind

MARK A. SABBAGH AND LINDSAY C. BOWMAN

INTRODUCTION: DEFINITION AND DISTINCTIONS

Everyday experience of the social world presents a cognitive puzzle. Although the actions that others take and the outcomes that those actions bring about are objectively clear, the reasons for those actions are not. There are many possible reasons for human actions, but the most proximal ones concern actors' mental states: their intentions, beliefs, and desires. Our conceptual understanding that mental states cause observable behavior, and our understanding of how they do so, is referred to as a "theory of mind" (e.g., Wellman, 1990). The term "theory" is used to capture the fact that although we cannot see others' mental states, we hypothesize their existence and make probabilistic judgments about their specific contents based on a range of relevant evidence. The adult theory of mind is also representational in the sense that we understand that the particular mental states that others hold are constrained by, but not copies of, some true state of affairs (Perner, 1991). A representational theory of mind is the foundation of our adult-like understandings of subjectivity—that two individuals can have different desires, beliefs, intentions, or even interpretations of some particular situation (Carpendale & Chandler, 1996).

The importance of a representational theory of mind is woven into every aspect of social experience in which making sense of others' actions is paramount. For instance, a jury's ability to discern the extent to which a particular crime was intended affects whether it is willing to hold the person responsible for the act and also the severity of the penalty jury members feel is appropriate to impose. (See Kaplan, 2001.) In dramatic arts, the intrigue and suspense of tragic and comic stories alike rely on the audience's abilities to track the subjective and varying mental states of the different characters as they develop within the story. As an example, consider the denouement of *Romeo and Juliet*. On their surface, the events that unfold make little sense—Romeo appears where he was told to look for Juliet, sees her there sleeping, drinks poison, and dies. But our theory of mind allows us to make sense of the tragedy—although we in the audience know that Juliet has taken a potion that has caused her to sleep, the forgoing events prevented Romeo from learning this information and so he falsely believes that she is dead. This situation makes sense only because of our abilities to reason about the hidden internal mental states that motivate human action.

The goal of this chapter is to provide a wide, albeit necessarily incomplete, survey of the current empirical and theoretical literature investigating the origins and development of theory of mind understandings. Our approach follows the historical contours of that literature while also highlighting what we believe are important new directions for work in the field along the way. We now turn to distinctions that we believe are important to draw in order to set the stage for our review.

Theory-Based Versus Immediate Induction

The example of *Romeo and Juliet* is an important one because it also underscores a distinction in theory of mind reasoning that sometimes is ignored—the distinction between theory-based versus immediate induction about the contents of others' mental states (c.f., Hughes & Leekam, 2004; Sabbagh, 2004). Theory of mind inferences based on theory-based induction occur when we make a probabilistic inference about the contents of another person's mental states based on our fundamentally conceptual, theoretical knowledge of how mental states arise from their underlying causes. In Romeo's case, our understanding of his false belief about Juliet's status comes from our theoretical understandings of how Romeo's idiosyncratic experiences would have shaped his mental representation of the situation (that he believes Juliet is dead) and caused his mental representation to be different from the audience's own (the audience knows she is only sleeping). Some of Romeo's experiences that could contribute to his belief about Juliet might have included hearing others talk of Juliet's death and her deathly appearance in the crypt. Importantly, we can make these inferences about Romeo's false belief and their origins without additional information from Romeo himself. Without Romeo saying a word or shedding a tear, our theory-based inductive inferences can render understandings of Romeo's likely (false) beliefs about Juliet's status.

In contrast to theory-based induction, theory of mind judgments that are based on immediate induction occur when we make inferences about others' mental states that are based on information that is available in the immediate situation. Perhaps the paradigm example of this immediate mental state induction is emotion recognition. For the most part, one does not need to know anything special about a person and his or her idiosyncratic histories to make a reasonable judgment about the person's emotional state; instead, these inferences are made on the basis of apparent information, such as facial expression, tone of voice, body posture, and the like. By "immediate," we do not mean to claim that the inferences are automatic or obligatory. (See Apperly, Riggs, Simpson, Chiavarino, & Samson, 2006.) Indeed, it seems likely that the inferences are still probabilistic ones insofar as there is no direct mapping from a particular expression to a particular state, especially when the information is partial or ambiguous (e.g., Russell, 2003). Nonetheless, the ways in which this kind of inference differs from theory-based inferences has important implications for understanding the conceptual, cognitive, and perceptual mechanisms that are likely to be involved in theory of mind reasoning.

It is very much an open question in the field as to whether theory-based versus immediate induction both rely on some underlying theory of mind ability, perhaps each related to a basic motivation to understand others' behaviors in mentalistic terms. In recognition of this uncertainty, some researchers and theorists prefer to talk about the general project of making judgments about other's mental states as "mentalizing" or "mindreading" while reserving the term "theory of mind" for those more circumscribed, theory-based judgments (e.g., Apperly, 2010). Because the types of inferences are so different and thus likely have different cognitive and neurobiological underpinnings, we believe that it is important to maintain the distinction.

Importantly, although there is a substantial literature on the development of children's immediate inductive inferences about others' mental states (e.g., sensitivity to gaze direction, joint attention, some aspects of prosodic understanding), we focus our chapter on the literature tracking developmental changes in young children's more theory-based conceptual understandings.

Cold/Cognitive versus Hot/Affective Theory of Mind

A second, possibly related, distinction that is relevant in the literature concerns broad differences in the content of mental state inferences. In particular, Shamay-Tsoory, Aharon-Peretz, and Perry (2009) have argued that there are meaningful differences between the abilities required to reason about "cold" cognitive mental states (e.g., knowledge and belief) versus "hot" affective mental states (e.g., emotions). Of course, there is little doubt that making inferences about others' cognitive states has consequences that are different from thinking about emotional states. The argument for why this distinction might be related to the theory-based versus immediate distinction just discussed is because the tasks that are used to measure understandings of cold cognitive mental states typically are theory-based induction tasks (e.g., the false-belief task; Wimmer & Perner, 1983); whereas the tasks used to measure understandings of hot affective mental states typically are immediate induction tasks (e.g., the "Reading the Mind in the Eyes Task"; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). However, the two sets of distinctions are at least theoretically orthogonal. Making accurate inferences about an actor's likely emotional states sometimes requires theory-like understanding of how emotional states are expressed in contexts. For instance, say you see a friend receive a gift that you know does not suit him or her; the person may feign delight despite an underlying disappointment (Harris, Donnelly, Gus, & Pitt-Watson, 1986). In these cases, conceptual knowledge of the underlying causes of emotion and emotional expression are necessary to understand the reasons for the particular expression in the moment. Likewise, some cold-cognitive states can be readily judged from immediately available information. For instance, whether people are ignorant about a particular topic can be apparent from their hesitant tone of voice, their closed body posture, and other kinds of information that can be gleaned in the immediate surround. (See, e.g., Birch, Akmal, & Frampton, 2010.) These examples illustrate the ways in which the hot–cold distinction is separable from the distinction between theory-based and immediate inferences about mental states. In this chapter, we focus primarily on work that covers theory-based inferences about both cold cognitive and hot affective mental states.

DEVELOPMENTAL TRAJECTORY OF THEORY OF MIND

False-Belief Development

Throughout the 1980s and 1990s, research in theory of mind was dominated largely by studies of young children's explicit understanding of false belief—beliefs that for some reason do not comport with some true state of affairs. The false-belief task comes in two basic varieties. In the "location change" false-belief task (e.g., Baron-Cohen, Leslie, & Frith, 1985), children hear a story often acted out in a puppet show, in which a protagonist (e.g., Sally) hides a toy in one location and then leaves the scene. In her absence, a second character (e.g., Anne) moves the toy to a new hiding place. Thus, Sally now has a false belief about the location of her toy; she believes it is where she left it, but really the toy is somewhere else. For the test question, children are asked where Sally thinks her toy is. In the contents-change task (e.g., Gopnik & Astington, 1988), children are shown a box that is known to have familiar contents, such as a box of candy. Children are then shown that the box in fact contains something other than candy, such as pencils. Here children are asked what a puppet (e.g., Snoopy), who has never seen inside the box, will think is inside. The idea is that children must recognize that Snoopy should look at the box and then have a false belief that the box has candy, because he has no reason to think otherwise. These tasks assess children's explicit understanding of false beliefs because they require children to provide conscious, declarative responses (either verbal answers or nonverbal points) to test questions asked by an experimenter.

False-belief research and the use of these types of tasks saw relatively intense focus for two reasons. The first is theoretical. A key feature of an adult-like understanding of mental states is that they are "representational"—they are subjective, person-specific representations of some external reality that are both based on and constrained by idiosyncratic experience. This adult-like understanding of mental states allows for the recognition that a given person's mental states can be different from another's or different from the reality that the mental states are supposed to represent. The false-belief tasks provide an elegant and relatively simple test of this complex adult-like understanding by asking children to recognize that a person will act in a way that does not comport with reality because of how that person is mentally representing the world. For example in the location-change false belief task, children are asked to report that Sally will look where she falsely believes the toy to be (i.e., where she left it) rather than where children know it really is. Thus, the logic of the false-belief task as a diagnostic for this representational theory of mind understanding generally is considered to be compelling and clear. (See Perner, 1991.)

The second and perhaps more important reason for the focus on false belief is that the task seemed to yield surprising results that demanded further investigation. The false-belief task normally is given to preschool-age children from 3 to 5 years old, and this period sees rapid development in children's performance. By 5 years old, children typically respond in line with an adult-like theory of mind. But more intriguing is 3-year-olds' tendency to fail the task systematically—instead of answering randomly (which would be expected if they were merely confused), 3-year-old children make consistent errors. In the location-change task, they consistently report that Sally will look where the toy is truly located rather than where she left it and thus should believe it is. In the contents-change task, 3-year-olds report that Snoopy will think that there are pencils in the candy box, even though he has no reason to believe that. These results with 3-year-olds were surprising to many in part because of the simplicity of the task and how intuitive correct responses in the task appear to be to adults. Three-year-olds' poor performance on these simple tasks also seemed inconsistent with their abilities to negotiate more naturalistic scenarios in which false-belief understanding might be required (e.g., Reddy, 1991, 2007). For instance, Bartsch and Wellman (1995; also Shatz, Wellman, & Silber, 1983) noted that at least some children even younger than 3 years old can make contrasts between the contents of a mental state and reality in their natural language (e.g., in conversations, 3-year-olds may say things like "You think he's sleeping but really he's awake"). Thus, many investigations focused heavily on the question of why 3-year-olds were systematically failing the task, and what might be done to improve their performance.

A leading hypothesis for why 3-year-olds consistently fail the false-belief task was (and to some extent, still is) that the task might be difficult for reasons besides the requirement to think about false beliefs. Specifically, the standard task has several events to keep track of, requires a modicum of linguistic understanding, and pragmatic understanding about test questions (e.g., Bloom & German, 2000; Siegel & Beattie, 1991; Westra, 2016). Wellman, Cross, and Watson (2001) summarized the attempts of a range of studies conducted between 1983 and 1998, many of which were aimed at simplifying the false-belief task to see whether 3-year-olds might show performance that better matched with the naturalistic data. Among the variables identified in the meta-analysis were changes to the test question, the nature of the objects that were involved, the extent to which children themselves were involved in the setup of the task, the length of the narrative, and the extent to which the task involved familiar real-world situations (e.g., deception). The meta-analysis showed that some of these modifications to the task were successful in leading children away from the "reality error"—in the change-location task, they no longer had a systematic tendency to report that the protagonist would look where the object really was. Yet there was no evidence that task modifications led to systematically correct performance in a group of children younger than around 4 years of age. It was concluded that there is a robust age effect that is not easily explained by any particular feature of the task. Rather, the authors concluded that between the ages of 3 and 5, false-belief performance improves because, during that time, children acquire the conceptual understandings that are necessary to make explicit decisions about others' representational mental states.

The shift that occurs in 3- to 5-year-old children's false-belief reasoning also appears to be stable across cultures. One convincing demonstration of this global pattern comes from a study by Callaghan et al. (2005) that used precisely the same methods with preschool-age children from Canada, Samoa, Peru, and India. Despite the diverse backgrounds, all groups showed evidence of a shift between the ages of 3.5 to 4.5 years old. A similar pattern emerged in a meta-analysis that focused on the comparison of Chinese and North American children, two regions where there has been substantial false-belief research (Liu et al., 2008). Despite many potentially meaningful cultural and linguistic differences between Chinese and North American children, the developmental timetable of false-belief reasoning is remarkably similar. (See also Sabbagh, Xu, Carlson, Moses, & Lee, 2006.)

Summary

Research using the false-belief task has provided substance to the claim that children's theory of mind understanding goes through an important transition over the preschool years that is relatively stable across cultures and not strongly affected by the specific way in which it is tested. Accordingly, it is now common to treat the false-belief task and other related tasks—such as appearance-reality tasks (Flavell, Flavell, & Green, 1983; see, e.g., Rakoczy, Fizke, Bergfeld, & Schwarz, 2015 for related skills)—as a reliable measure of preschoolers' social cognitive understandings. This reliability and validity of the false-belief task has enabled researchers to probe the association between theory of mind and social competencies more broadly construed. The false-belief task also has provided an avenue for investigating the experiential and neurobiological factors that are associated with theory of mind and its development. We discuss both of these directions of research in separate sections later in the text.

Yet it is important to note that the question of whether young preschoolers' failures on the false-belief task are truly attributable to conceptual deficits—that is, whether they are attributable to a misunderstanding of the connection between beliefs and the reality that they are supposed to represent—is not settled. Similarly, the question of whether preschoolers' changing performance on false-belief tasks is attributable to conceptual changes also remains unsettled. The controversy has resurfaced largely over the last decade due to findings from false-belief studies with infants, which we turn our attention to next.

Implicit False Belief: New Debates and Controversies

The distinction between implicit and explicit understandings is a complicated one that often is defined more in terms of the response modalities than in terms of the underlying character of the knowledge represented (Dienes & Perner, 1999). A task is considered to be explicit when responses are declarative and are taken to result from decision-making processes in response to some test question. Explicit responses can be verbal or nonverbal (labeling or pointing) but generally are communicative and intended to be answers to the experimenter's questions. An implicit task reflects nondeclarative, obligatory natural responses to ongoing scenes. For example, an implicit task might monitor gaze to look for evidence of either surprise (when some event violates an expectation) or predictive looking, either of which can be based on some putative conceptual understanding that is isolated by the experimental design. The crucial difference between these response modalities is that explicit tasks are thought to require not just the conceptual understandings but also the cognitive resources that are necessary for decision making in the moment. Implicit tasks do not make these same additional demands. In the present context, implicit tasks provide an opportunity to reopen the question about whether children fail "explicit" false-belief tasks because they lack the conceptual understandings or because they lack the sufficient domain-general cognitive resources to negotiate the complications of the task.

There is now ample evidence suggesting that young infants accurately predict others' behaviors in false-belief-type scenarios when their predictions are assessed more implicitly (for reviews, see, e.g., Baillargeon, Scott, & He, 2010; Heyes, 2014), potentially undermining the conclusion that 3-year-olds fail false-belief tasks due to conceptual deficits. A pioneering study by Onishi and Baillargeon (2005) used the violation-of-expectation methodology in which 15-month old infants were familiarized with a scene where an agent hid an object in one of two locations. The object then moved to an alternative location, either while the agent was observing (true-belief condition), or while she was unable to see the object's movements (false-belief condition). Looking times were measured while infants observed the agent reaching to either the outdated or current object location. In false-belief conditions, infants showed evidence of surprise when the agent searched where the object really was, thereby suggesting that they expected the actor to act in accordance with their false belief. Similar findings have been reported with younger children, (e.g., Kovács, Téglás, & Endress, 2010; Surian, Caldi, & Sperber, 2007) and with infants in implicit paradigms other than violation of expectation, such as predictive looking (e.g., Southgate, Senju, & Csibra, 2007; Surian & Geraci, 2012). Positive findings also have been reported for implicit versions of contents-change false-belief tasks (e.g., Scott, & Baillargeon, 2009).

The strong interpretation of these findings is that even young infants have some very early emerging understanding—if not innate understanding (e.g., Leslie, Friedman, & German, 2004)—of false belief that is masked in studies that require an explicit response because of the peripheral cognitive demands required to negotiate this explicit response. To date, there is no obvious reason to doubt that implicit measures can show that infants make accurate predictions in false-belief scenarios. Nonetheless, there is considerable debate over how to best resolve the puzzle of why young infants appear to show competence with false-belief reasoning in implicit tasks while 3-year-olds do not in explicit tasks. We return to this question where appropriate throughout the rest of this chapter, as many sources of evidence can bear on the question. Broadly, for those who do not subscribe to the strong interpretation of the infant data, it remains possible that the mechanisms underlying performance on the infant false-belief tasks are fundamentally different from those that underlie preschoolers' performance. There have been two general approaches to characterizing how those mechanisms might differ: One focuses on domain-general accounts of infants' false-belief performance, and another posits two distinct systems for making judgments about others' mental states. We discuss each next in turn.

Domain-General Accounts of Infant False Belief

One argument for how to reconcile the infant–preschool false-belief discrepancy is that infants' responses in the violation-of-expectation or predictive looking tasks might be attributable to the dynamics of domain-general cognitive processing (e.g., Heyes, 2014; Sabbagh, Benson, & Kuhlmeier, 2013). For instance, Heyes (2014) reviewed the findings of 20 separate studies purporting to demonstrate false-belief reasoning in infancy and demonstrates how each might be better explained by infants' attention being captured by contextual novelty. Others (e.g., Ruffman, 2014) have argued that processes associated with episodic encoding might lead children to encode generalized behavioral rules about what typically happens in everyday situations (i.e., "People typically look for things where they last left them"). Of course, much more research needs to be done to determine whether these domain-general mechanisms can provide a full accounting of infants' performance in implicit false-belief tasks. Nonetheless, what is notable about these positions is their attempt to account for performance in the infant paradigms without reference to psychological understandings whatsoever. If these positions are proven true, then there may be little continuity between infants' understanding and the developments that occur during the preschool years.

Dual-System Account

A second possibility for the discrepancy between infant and preschool false-belief performance is that infants are performing a kind of psychological reasoning, but it is one that is limited to a relatively circumscribed set of situations. This possibility is sometimes called the dual-systems account. (See, e.g., Apperly & Butterfill, 2009; Low, Apperly, Butterfill, & Rakoczy, 2016.) The two systems proposed include an early emerging one that registers "belief-like states" in a fast, efficient, but stereotyped manner and a second, later-emerging one that computes beliefs per se but does so more slowly and requires domain-general cognitive resources that are necessary to maintain a changing, flexible representation of others' minds. The specific claim is that infants' looking patterns in the implicit false-belief tasks may be attributable to the fast system while preschoolers may be relying on the slower one to make their explicit responses. This general model has catalyzed a small but burgeoning body of research aimed at better understanding the different characteristics of the fast and slow systems, and particularly the signature limitations of the fast system (see e.g., Low & Watts, 2013; Schneider & Low, 2016) and the cognitive costs of the slow system (e.g., Apperly, Back, Samson, & France, 2008). What is important, however, is that this model conceptualizes infants' performance in the false-belief task as relying on fundamentally mentalistic concepts. Because both systems are fundamentally mentalistic, it is possible that there are developmental continuities between infants' and older children's theory of mind understandings, such that early understandings may be predictive of later ones. (See, e.g., Thoermer, Woodward, Sodian, Perst, & Kristen, 2013).

Summary

In laboratory tasks that rely on implicit responding, infants show evidence that they expect people to act in accordance with false beliefs. More work needs to be done to characterize both whether and how the cognitive underpinnings of infants' performance in implicit false-belief tasks differs from those supporting preschoolers' performance in explicit false-belief tasks. We see this as an exciting area for research in the field and one that recognizes the remarkable findings that continue to appear in the infant literature, while still acknowledging the importance of developmental achievements over the preschool years that have been demonstrated by years of careful research.

Beyond False Belief: Understanding Other Mental States

Theory of mind understanding is not limited to an understanding of false beliefs. The ability to make sense of and predict others' behavior involves understanding multiple interconnected mental states. To take a simple example, we can understand the action of going to the refrigerator for milk as a behavior that is motivated by a belief that the milk is in the refrigerator and a desire for milk; we would predict a different behavior if we knew someone thought there was milk in the refrigerator but did not desire it. The role of multiple mental states becomes even more apparent when we consider not just simple actions with objects but interactions with others. Our ability to successfully, strategically collaborate, negotiate, or compete with others is best served when we are able to give full consideration to others' intentions, desires, knowledge, and beliefs (Tomasello, 2014). Although the cognitive and conceptual mechanisms underlying our abilities to reason about intentions, desires, and knowledge may be different from those underlying an understanding of beliefs and false beliefs, an understanding of these additional mental states is nonetheless critical to our everyday social reasoning (Wellman, 2014).

This broader view of theory of mind suggests that its developmental trajectory has its origins in early social perception. By 3 months of age, infants follow another's eye gaze (e.g., D'Entremont, Hains, & Muir, 1997; Moore, 1999) and engage in joint visual attention in social contexts (e.g., Corkum & Moore, 1998; Deák, Flom, & Pick, 2000). These abilities demonstrate a sensitivity to others' attention and to the perceptual information available to others, which may in turn be important for constructing an understanding of the connections between the observable world and internal mental states.

In this section we cover how, over the first 2 years of life, infants' understanding advances from initial sensitivities for noticing links between action and intention, to more complex conceptions of person-specific desires, and to an understanding of the connections between prior experience and states of knowledge/ignorance. These conceptual advances are part of the larger developmental framework for an adult-like representational theory of mind.

Understanding Intentions

By as young as 6 months old, infants already seem to possess some understanding that others' actions are motivated by underlying intentions and goals. For example, after habituating to an agent's hand reaching for and grasping a teddy bear on the left side of a stage, 6- and 9-month-olds look longer when the hand keeps the same path of motion (reaching to the left) but demonstrates a distinct goal (grasping a ball instead of a teddy bear; Woodward, 1998). These results suggest that infants encode others' actions in terms of their intended goal and not simply in terms of specific paths of motion. The ability to encode actions as goal-directed is also evident in infants' imitative behavior: 7-month-olds selectively imitate an experimenter's actions when the experimenter has a clear, intended goal (e.g., grasp an object) but do not imitate actions that do not have clear goals (e.g., touch object with back of hand; Hamlin, Hallinan, & Woodward, 2008; Thoermer et al., 2013).

By 10 to 12 months, infant understanding of intentions begins to become separate from observable actions such that they can impute the intentions behind failed actions as well (Brandone & Wellman, 2009). For example, after habituating to a scene in which an agent reached in an upward arc over a tall barrier to almost touch but never fully retrieve a ball, infants looked longer when viewing the agent reach in a similar upward arcing path of motion when the barrier was removed than when viewing a direct reach toward the ball (which was now possible given the lack of barrier). This looking pattern suggests that by the end of the first year of life, infants are able to encode others' actions in terms of their underlying goals (e.g., to retrieve a ball), even when the goal itself is not fully demonstrated.

A substantial literature has built up around specifying the conditions under which children appear to impute intentions and goals to actors in simple situations. (See, e.g., Robson & Kuhlmeier, 2016, for a recent review.) Within this literature, perhaps one of the most intriguing empirical questions concerns whether the action or goal that is being observed is one that also can fit within infants' own action repertoires. For instance, infants who cannot themselves grab objects also appear to have difficulty attributing grasping intentions to others (e.g., Falck-Yitter, Gredebäck, & von Hofsten, 2006; Kanakogi & Itakura, 2011). Intriguingly, however, children's abilities to attribute intentionality to an action can change rapidly with their own experience. In a series of remarkable studies by Sommerville and colleagues (e.g., Sommerville, Woodward, & Needham, 2005), 3-month-olds who were unable to grasp objects were given experience with Velcro mittens that allowed them to "grasp" Velcro objects around them. Although the infants did not show evidence of intention attribution prior to this experience, they did after the experience when shown an actor wearing those same gloves. These findings, and others like them, show that from very early, infants have the capacity to interpret others' actions as intentional, but that those attributions are constrained to intentions that are within the infants' own action repertoire. By the end of the first year, infants come to take a broader, criterion-based view of what kinds of actions (and from what kinds of agents) are intentional (e.g., Shimizu & Johnson, 2004).

Understanding Desires and the Subjectivity of Mental States

At the end of the first year of life, infants also begin to understand actions in terms of their underlying desires. After viewing an agent express desire for a specific toy (e.g., smiling and happy speech toward a particular toy), 12-month-olds look longer when the actor picked up a different toy for which the agent expressed no previous desire, suggesting infants expect others' actions to be based in desires (Phillips, Wellman, & Spelke, 2002). A few months into the second year of life, infants' understanding of intentions and desires incorporates an understanding that mental states are person-specific and subjective. For instance, by 12 to 14 months, infants will follow an agent's gaze around a barrier, even if this requires the infant to lean or move behind the barrier, suggesting the infant understands that the agent may see something they themselves do not see (Moll & Tomasello, 2004). By 18 months, infants can apply this understanding of subjective experience to others' desires as well: When given a choice to offer an experimenter one of two snacks, 18-month-olds will offer the experimenter the snack for which she previously expressed desire (i.e., said, "Mmm, yummy!" after eating broccoli), even if the experimenter's desire contrasted with the infants' desire for the alternate snack (i.e., infant prefers crackers over broccoli) (Repacholi & Gopnik, 1997). These findings demonstrate that toddlers can recognize that desires are subjective and distinct; that they are specific to individuals and can sometimes contrast with the infants' own. What is perhaps most intriguing about these findings is that the understanding that two individuals can have different desires regarding some particular state of affairs emerges well in advance of the understanding that two individuals can have different beliefs (e.g., Wellman & Liu, 2004). Findings such as these help focus our understanding of what specifically develops over the preschool years—namely, it is not the understanding of subjectivity per se but rather the understanding of how subjectivity applies to a range of mental states.

Understanding Knowledge and Ignorance

There is also evidence to suggest that infants understand something about the mental states of knowledge and ignorance. Twelve-month-olds notice when an experimenter's prior experience would make her knowledgeable or ignorant about the location of an object and are more likely to point for the experimenter's benefit when she is ignorant (Liszkowski, Carpenter, & Tomasello, 2008). Twelve- and 18-month-old infants also can correctly identify an object whose presence an experimenter was previously ignorant of (Tomasello & Haberl, 2003). After playing together with two specific objects, when a third object was introduced to the infant in the experimenter's absence, the infant gave the experimenter the novel toy when the experimenter returned and exclaimed, "Wow! Cool! Can you give it to me?" These findings reveal that around children's first birthday, they are sensitive to the factors that can make a particular item or situation novel from one person's perspective even if it is not novel from their own perspective. Thus, similar to their understanding of desires, by the end of the first year of life, infants show some nascent appreciation of the person-specific nature of mental states.

Developmental Progression from Desire to Belief Understanding

As was apparent in the foregoing discussion, there is now strong evidence of a developmental progression from an early understanding of desires to a later explicit understanding of beliefs and thoughts. For some time, the progression was difficult to pin down experimentally because the tasks used to measure desire and belief understanding were different in a number of ways. Additional evidence comes from natural language; although children use desire words (e.g., "want," "like") and can predict that storybook characters will act according to their desires by around age 2 years, it is not until around 3 years of age that they use belief words (e.g., "think," "know") and consistently make predictions for characters based on beliefs (Bartsch & Wellman, 1995). Moreover, the desire–belief progression holds across tasks that are carefully matched on procedural methodology, linguistic structure, and materials (e.g., Wellman & Liu, 2004).

One outcome of this research has been the construction of a theory of mind scale that can assess children's achievements of various conceptually relevant milestones that emerge prior to false-belief understanding (Wellman & Liu, 2004). In addition to its growing use as an assessment tool for research purposes, research using the scale has shown that the order in which children acquire the milestones tends to be fairly stereotyped within a given population. However, there are some relevant cross-cultural variations in the order in which milestones are acquired (e.g., Shahaeian, Peterson, Slaughter, & Wellman, 2011; Wellman, Fang, Liu, Zhu, & Liu, 2006). In brief, children growing up in cultures emphasizing collective knowledge and dispute avoidance, such as China (Wellman et al., 2006) and Iran (Shahaeian et al., 2011), understand knowledge and ignorance in others before they understand the person-specific nature of beliefs; whereas children growing up in cultures emphasizing individuality and diverse opinions, such as the United States (Wellman & Liu, 2004), Australia (Shahaeian et al., 2011), and Germany (Kristen, Thoermer, Hofer, Aschersleben, & Sodian, 2006, as cited in Wellman, 2014), understand the person-specific nature of beliefs prior to understanding knowledge and ignorance. However, the progression of understanding desires before beliefs appears especially robust: It is consistently evident across each of these countries and cultures.

An intriguing question concerns why an understanding of desires might develop prior to an explicit understanding of beliefs. One possibility is that beliefs are especially difficult; they require representing not just the contents of someone else's mental states but also the relationship that those contents are supposed to bear with respect to some true state of affairs (Perner, 1991; Wellman, 2002). For example, someone can believe he left a bag of apples in the car, while someone else can believe the apples were left on the kitchen counter. At least one of the individuals' beliefs will be wrong because there is some true state of affairs that the belief is supposed to represent. Desires, in contrast, do not have the same requirement; one person can like broccoli while another likes Goldfish crackers; yet no one desire is right or wrong. Perner (1991) expressed the difference as one of "thinking of" versus "thinking that." One can think *of* oneself being on a beach without causing any representational conflict. However, if one were to think *that* one was on a beach when in fact one was in a dreary office, this would indeed be cause for concern. Some researchers have called this additional requirement metarepresentation because it involves thinking not just about the contents of a mental state but also about the specific way in which those contents are supposed to represent reality. (See Perner, 1991.) Belief-reasoning may develop later than desire-reasoning because additional cognitive or computational capacities may be required to understand the metarepresentational nature of beliefs, and these capacities may not exist until later in development.

Summary

Theory of mind development begins well before the emergence of explicit false-belief understanding. Children's early understanding of intentions, desires, and knowledge are sophisticated insofar as they incorporate an understanding that the contents of these mental states can differ across people, relative to their specific experiences and proclivities. There is a specific developmental progression by which these understandings emerge from infancy through the preschool period—understandings of motivational mental states, such as desires and intentions, emerge and become more sophisticated prior to equivalent understandings of epistemic mental states, such as knowledge and beliefs. More work is necessary to better understand why this developmental progression exists, but there are reasons to think that epistemic mental states (beliefs in particular) have unique conceptual complexity.

Theory of Mind Development in Middle and Late Childhood

Just as theory of mind reasoning does not begin with false-belief understanding, so too does it not end there. Performance on standard false-belief tasks (e.g., the location-change and contents-change tasks described earlier in the section titled "False-Belief Development") typically reaches ceiling shortly after 5 years old (e.g., Wellman et al., 2001). Perhaps because of this methodological issue, investigation of theory of mind in school-age children and adolescents is still quite limited. Some researchers have used second-order false-belief tasks in which participants are required to reason not just about a person's belief but about a person's belief about someone else's belief or other mental state (e.g., Sally thinks that Anne thinks that the toy is in the basket; John thinks that Mary is happy to see him; Perner & Wimmer, 1985). Most children fail these second-order tasks prior to 7 years of age (Miller, 2009). Third- and fourth-order false-belief tasks (e.g., Sally thinks that Anne thinks that Mary is in the kitchen) are even more challenging. The increasing difficulty of the task demands can yield individual differences in performance into late childhood and adolescence (Liddle & Nettle, 2006).

Critically, though, the achievement of belief and false-belief reasoning, even as assessed in more challenging tasks, does not equip children with all that is necessary to understand the nuanced connections between minds and everyday behavior. To accurately predict and make sense of human action and interaction, it is important to also understand how different mental states interact and how mental states can be affected by time and context. Assessments of these richer, more flexible mental state understandings reveal developmental advancements extending beyond preschool, into middle and late childhood and adolescence. (See Lagattuta et al., 2015, for review).

Understanding Interactions Between Mental States

Children's understanding of how different mental states interact with each other to affect behavior emerges after children enter school. For example, Sayfan and Lagattuta (2008) asked preschoolers, school-age children, and adults to predict and explain the emotions of infant, child, and adult protagonists in threatening situations (e.g., encountering a snake). Accurate responses required taking into account the amount of knowledge that each type of protagonist would have about the situation. Adults and school-age children, but not preschoolers, were able to consider that more knowledgeable protagonists (e.g., adults) who understood the threat would be more fearful than less knowledgeable protagonists (e.g., infants and children) who may not know that a given situation is threatening.

"Interpretive" theory of mind tasks also require participants to reason about others' mental states based on their idiosyncratic knowledge (e.g., Carpendale & Chandler, 1996). In these tasks, participants are asked to predict how different individuals will interpret a picture that is partially occluded based on the individuals' prior experience and knowledge of the image (e.g., What will Sally and Anne each think this is a picture of, given Sally never saw the unoccluded picture but Anne did?). These tasks are particularly challenging because they require participants to construct different beliefs for different individuals. Children typically do not pass these tasks until 6 to 7 years old, and performance continues to improve into adulthood (Lagattuta, Sayfan, & Blattman, 2010; Lagattuta, Sayfan, & Harvey, 2014; Ross, Recchia, & Carpendale, 2005).

Over middle childhood, children also come to understand the influence of people's thoughts on their emotions. Between 5 and 10 years of age, children progress in their understanding that even though two people might feel the same way about a given negative situation, subsequent thoughts can affect the intensity of those feelings. For instance, 8- and 10-year-old children accurately judge that two children who break their arms will be sad, but the one who sees a bright side (e.g., "all my friends can sign my cast") will feel less sad than someone who dwells on the negative aspects (e.g., "my arm will be itchy"); whereas 6-year-old children judge the two characters' feelings as similar despite their different views (Bamford & Lagattuta, 2012). These findings nicely illustrate ways in which children's theory of mind understandings per se becomes increasingly mature over middle childhood.

Mental State Reasoning in Context

Older children also begin to consider aspects of the current context that can inform judgments about the underlying intentionality of people's actions. One nice example of this is the faux pas task in which children are told stories in which a character accidentally says something to another character that is potentially offensive or mean given what children know about the context (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999). For instance, in one story, a woman just moved into a new house and bought new curtains. Her friend visited later that day for the first time and insults the curtains, saying "Oh, those curtains are horrible! I hope you are going to get new ones!" The ability to understand both that the utterance would be offensive and that the offender did not intend harm develops over 5 to 11 years old (Banerjee & Watling, 2005; Baron-Cohen et al., 1999).

Summary

Much of the complexity of human interaction relies on sophisticated theory of mind understandings that are built on the insights that are achieved during the preschool years but also go well beyond them. In particular, during the school years, children come increasingly to appreciate the intertwined nature of mental states and how they are affected by context. Surely more aspects of a sophisticated adult-like theory of mind continue to develop into adolescence. (See, e.g., Boyes & Chandler, 1992.) An important direction for future research in this area is delineating those areas and establishing through individual differences studies whether they make important contributions to children's abilities to successfully navigate the increasingly complex social worlds that they inhabit as they get older.

FACTORS ASSOCIATED WITH THEORY OF MIND DEVELOPMENT

For the remainder of the chapter, we are going to describe work that details what is known about the cognitive, neurobiological, and experiential factors that affect theory of mind development. Much of this work is focused on the transitions that occur during the preschool years and thus on the factors that affect false-belief understanding. Thus, we return our attention to that age group and revisit some of the issues that arose during our prior discussion. We do ask, however, that the reader bear in mind that although the most is known about the false-belief task, we believe that a critical direction for future research is understanding more about the factors that affect the developments that occur both before the false-belief milestone and afterward.

Executive Functioning and Theory of Mind

"Executive functioning" is the term used to refer to the broad suite of interrelated domain-general cognitive skills that are necessary for goal-directed action, including working memory, set shifting, and inhibitory control (Garon, Bryson, & Smith, 2008; Zelazo, Carlson, & Kesek, 2008). There are several reasons to believe that executive functioning skills would be important for correct performance on theory of mind tasks and false-belief tasks in particular (Frye, Zelazo, & Palfai, 1995; Hughes, 1998; Russell, 1996). For instance, in the location-change false-belief task, children have to watch a story in which a character returns to a scene and is looking for an object that has moved in her absence. Children's natural tendency might be simply to tell the story character where her toy is, and so inhibitory control is necessary to overcome this prepotent tendency and answer the test question. Similarly, it is natural and habitual to tell others (by pointing) where objects truly are, and so inhibitory control may be required to make the noncanonical response to point to a location where something is not (Carlson, Moses, & Hix, 1998). In a separate analysis, Frye and colleagues (1995) suggested that the structure of a false-belief task is similar to a set shifting task like the dimensional-change card sorting task.

Two lines of evidence show that executive functioning is associated with false-belief reasoning. The first and more extensive line takes an individual differences approach in which children's performance on a battery of tasks that measure executive functioning is associated with performance on a false-belief battery (e.g., Carlson & Moses, 2001; Hughes, 1998). This work consistently has shown that executive functioning skills are moderately correlated with false-belief performance. (See Devine & Hughes, 2014, for a meta-analysis.) The correlation between executive functioning and false belief is typically higher for response-conflict executive functioning tasks (e.g., Stroop-like tasks) that require children to choose between one of two competing responses based on a novel task rule than it is for delay tasks that simply ask children to withhold a response for some time (Carlson & Moses, 2001). Some have suggested that this preferential relation is because response-conflict tasks combine component executive functioning skills in much the same way that false-belief tasks do (Hala, Hug, & Henderson, 2003).

The second line of evidence for links between theory of mind and executive functioning comes from research showing that experimentally titrating the executive demands of the false-belief task has predictable effects on children's performance. For instance, children's performance on the location-change false-belief task improves if the object is moved to a location away from the immediate scene (Wellman et al., 2001), presumably because taking the object away weakens children's prepotent tendency to inform the character about the true location of the object. Similarly, other studies saw improved performance in preschoolers by asking children to respond to the test question in a novel way (e.g., using a cardboard arrow), which likely reduced the need for children to overcome the prepotent tendency to point to where something really is (Carlson et al., 1998; Coulliard & Woodward, 1999). Likewise, even older children's performance can be worsened by increasing the executive demands of the task. Careful work by Friedman and colleagues (e.g., Friedman & Leslie, 2004) showed that adding another location worsens children's false-belief performance, likely by increasing the complexity of the decision they have to make. The same is true for older children; theory of mind reasoning is negatively affected by task modifications that increase the load on executive functioning (Qureshi, Apperly, & Samson, 2010).

It is clear that false-belief tasks have executive functioning demands, and so having a modicum of executive functioning is certainly necessary for children to demonstrate false-belief understanding. Yet, from a developmental perspective, there is burgeoning evidence that although executive functioning skills may be necessary, they are *not sufficient* for preschoolers' false-belief reasoning. One line of evidence comes from a comparison of Asian and North American preschoolers. Preschool-aged children in Asian cultures typically show advanced executive functioning skills relative to their North American counterparts; for example, in one study by Sabbagh et al. (2006), 3.5-year-old Chinese children performed as well on a battery of executive functioning tasks as 4-year-old North American children. Yet, despite this substantial advantage in executive functioning, the Chinese preschoolers showed no parallel advantage in false-belief reasoning. The same pattern has been replicated in China (Tardif, Wellman, & Cheung, 2004) and with Korean children (Oh & Lewis, 2008). These findings and others (see, e.g., Benson & Sabbagh, 2009) demonstrate that although executive functioning skills may be necessary for false-belief reasoning, they are not themselves sufficient to account for individual differences in theory of mind, or for theory of mind developments.

It can be noted that these findings regarding the specific association between executive functioning and preschoolers' false belief development have implications for common interpretations of the infant implicit false belief findings. Recall that the findings with infants have led some to conclude that even young infants have the capacity for false-belief understanding but that the explicit tasks that are used with 3-year-olds mask those capacities because of their peripheral cognitive demands. The findings with executive functioning speak against this strong interpretation, at least as it pertains to the peripheral executive functioning demands of false-belief tasks; there are groups of children who have relatively advanced executive functioning skills yet still have difficulty with explicit false-belief tasks (see Low et al., 2016; Sabbagh et al., 2013).

Though executive functioning skills may not be sufficient to account for theory of mind development, they may nonetheless play an important role in supporting theory of mind development (see Moses, 2001; Moses & Sabbagh, 2007). Longitudinal studies have shown that individual differences in 3.5-year-old children's executive functioning skills are a unique longitudinal predictor of preschoolers' false-belief reasoning a year later (e.g., Hughes, 1998; Carlson et al., 1998; Flynn, 2007, see also Devine & Hughes, 2014). More recently, there is evidence that among 3.5-year-olds who fail false-belief tasks, individual differences in executive functioning predict the extent to which those children benefit from a training regimen designed to promote false-belief reasoning (Benson, Sabbagh, Carlson, & Zelazo, 2013). These findings show that in addition to being important for negotiating task demands, EF may play a causal role in children's acquisition of the conceptual understandings necessary for false-belief reasoning.

Summary

The findings detailing the association between theory of mind and executive functioning in preschool-aged children clarify both that false-belief tasks have executive functioning demands that affect children's performance, and also that immature executive functioning is not the only reason for 3-year-olds' failure at false-belief tasks. These findings support the view that conceptual developments that occur over the preschool years are critical to children's abilities to reason about the representational nature of mental states. There now also is evidence to suggest that executive functioning skills may play a causal role in the process of acquiring theory of mind concepts from experience.

Neural Bases of Theory of Mind Development

The cross-cultural synchrony in the broad trajectory of theory of mind development suggests that its broader parameters might be established by relatively specific neurobiological factors that also may change rapidly over the first 5 years of life. In adults, representational theory of mind understanding is associated with a distinct set of neural regions, including the right temporal–parietal junction (rTPJ), the precuneus, and the dorsomedial prefrontal cortex (DMPFC). (See Koster-Hale & Saxe, 2013, for a recent review.) There is now substantial evidence suggesting that the same is true for young children.

Functional Brain Development

A study using source-localized electroencephalogram (EEG) methods showed that the functional maturation of these areas that are associated with false-belief reasoning in adults, in particular the rTPJ and the DMPFC, was positively associated with performance on a theory of mind battery in young children (Sabbagh, Bowman, Evraire, & Ito, 2009). That is, children with more evidence of functional maturation in these regions showed better performance on the false-belief task and other similar tasks. A similar pattern of results has been shown using functional magnetic resonance imaging methods (e.g., see Gweon & Saxe, 2013). Although they are interesting in their own right, these findings also inform the debate about the nature of the developments that support theory of mind development over the preschool years. Recall that on the basis of the findings potentially showing false-belief competence during infancy, some have suggested that the transitions over the preschool years reflect the development of executive functioning skills rather than the development of theory of mind. The current findings provide evidence against this explanation of the developmental transitions in preschoolers by showing that it is the functional development of the theory of mind network per se that supports children's performance on false-belief tasks rather than areas that are important for domain-general cognitive skills like executive functioning.

The importance of the functional recruitment of the prefrontal cortex (PFC) in preschoolers' theory of mind reasoning has also been shown in a handful of event-related potential studies. (See Sabbagh, 2013, for a review.) For instance, Liu, Sabbagh, Gehring, and Wellman (2009) showed that 6-year-olds who could pass a relatively complicated false-belief task showed a left-lateralized frontal late slow wave effect as they responded to test questions. Intriguingly, children who were the same age but did not pass the tasks did not show that same effect, thereby suggesting that the recruitment of regions of the frontal lobe, perhaps especially the medial PFC (MPFC), is associated with good performance on theory of mind tasks. A frontal late slow wave effect has been associated with reasoning about beliefs in 6- to 8-year-old children (Meinhardt, Sodian, Thoermer, Döhnel, & Sommer, 2011) and in 7- to 8-year-old children (Bowman, Liu, Meltzoff, & Wellman, 2012), all using slightly different paradigms. Taken together, these findings converge on the fact that for young children, changes in theory of mind reasoning are associated with the functional maturation and recruitment of relatively discrete neural systems within the MPFC.

What is particularly intriguing about the MPFC regions is that they appear to be important for reasoning not only about beliefs but about desires as well. Work by Bowman and colleagues (2012) using event-related potentials showed that school-aged children's desire and belief reasoning were associated with mid-frontal activity, but only belief reasoning was associated with activity at right parietal sites, near the rTPJ. The association between the rTPJ and school-age children's reasoning about beliefs but not desires also was shown in a study that used functional near-infrared spectroscopy (Bowman, Kovelman, Hu, & Wellman, 2015), which directly measures blood flow to cortical regions. These findings shed light on the behavioral findings reviewed earlier that show that children exhibit sophisticated reasoning about desires before beliefs. In particular, the findings suggest that the developmental progression from desire to belief understanding may be paced in part by development of the rTPJ as it specializes to support an increasingly advanced understanding of beliefs. Converging support for this possibility comes from functional magnetic resonance imaging work by Saxe and colleagues (Gweon, Dodell-Feder, Bedny, & Saxe, 2012; Saxe, Whitfield-Gabrieli, Scholz, & Pelphrey, 2009), who showed that although 6-year-old children exhibit specialized activation in the DMPFC for reasoning about others' mental states, specialized activation in the rTPJ did not emerge until children were about 9 to 11 years old. The whole set of findings suggests that the MPFC regions that support theory of mind reasoning may reach functional maturity early in development and contribute to reasoning about a range of theory of mind skills. In contrast, the rTPJ, which appears to be especially important for reasoning about beliefs, has a longer developmental timeline that may begin in the preschool years but clearly extends well after.

Two noteworthy implications of these neurodevelopmental patterns suggest an early present role of the DMFPC in mental state reasoning more broadly and a protracted specialization of the rTPJ for belief reasoning specifically. The first is that the DMPFC may support a sort of general theory of mind that contributes to a host of theory of mind tasks and is involved whenever making judgments about others' mental states. The development of the functional specialization of the DMPFC also may represent a sort of rate-limiting factor such that the healthy development of these regions may have particularly profound effects on typical theory of mind reasoning. The second implication is that, echoing our review of the behavioral literature, the development of theory of mind understanding does not end when children become skilled at passing false-belief tasks (e.g., Dumontheil, Apperly, & Blakemore, 2010). Given the emergence of exciting new paradigms for studying advances in theory of mind reasoning that emerge after the preschool years, the stage is now set to better understand the behavioral correlates of these neurodevelopmental changes in the theory of mind network.

Dopamine

In addition to looking at functional and developmental changes in neuroanatomical structures, some recent work has investigated the extent to which neurochemical factors might be associated with theory-of-mind development. In this regard, particular attention has been paid to dopamine (DA) (e.g., Lackner, Bowman, & Sabbagh, 2010). DA is of particular interest because the region of the DMPFC that is important for theory of mind and implicated in its development lies along the mesocortical DA pathway. Moreover, DA signaling appears to be sensitive to a number of cognitive processes that have been thought to be important for developing correct inferences about others' mental states, such as adjusting predictions based on feedback. (See Sabbagh, 2016, for a recent review.)

There are now two pieces of evidence suggesting that individual differences in DA functioning may be associated with false-belief performance in preschoolers. The first is that individual differences on an indirect marker of DA functioning, spontaneous eye-blink rate, are associated with performance on a battery of false-belief tasks (Lackner et al., 2010). The second is that allelic variation in the D4 DA receptor gene (*DRD4*) is associated with theory of mind such that children with genes that promote more efficient DA signaling in the PFC show better performance (Lackner, Sabbagh, Hallinan, Liu, & Holden, 2012). The connection between theory of mind and DA was specific to the *DRD4* gene, which is most strongly predictive of DA binding in the PFC; similar associations were not seen for genes that are associated with either DA metabolism (i.e., *COMT*), or synaptic DA in the striatum (i.e., *DAT1*). These findings dovetail with the findings from the neuroanatomical studies in showing that healthy medial prefrontal development and functioning is associated with transitions in preschoolers' theory of mind understandings.

Broader Neural Network and Emerging Directions

Although the bulk of the work on the neurobiological bases of theory of mind has focused on the DMPFC and TPJ and their roles in children's conceptualizations of beliefs and desires, it is important to note that the network of regions implicated in mental state reasoning extends beyond these two regions. The superior temporal sulcus and the precuneus are part of a broader network of neural regions supporting theory of mind in adults, and these regions likely support aspects of theory of mind in development as well. For example, the superior temporal sulcus is recruited when school-age children infer intentions from action (Mosconi, Mack, McCarthy, & Pelphrey, 2005; Ohnishi et al., 2004) and when young adolescents think about the self (Pfeifer et al., 2009). The precuneus is recruited when children are required to reason about multiple mental states simultaneously (Gweon et al., 2012; Saxe et al., 2009). As noted, longitudinal work is necessary to establish a causal role for these aspects of brain development and the relevant behavioral changes in theory of mind reasoning.

Mu Rhythm and Neural Networks for Action Perception and Production

An exciting new direction in understanding the neurocognitive bases of theory of mind reasoning focuses on the neural processes that are specifically recruited for understanding and producing intentional action. In contrast to the work that looks at explicitly judging action as intentional or nonintentional or deciding how someone with particular intentions might act, this work examines the neural processes that are associated with the ongoing perception and execution of intentional actions. There are theoretical reasons that action and theory of mind development might be linked (Hunnius & Bekkering, 2014; Meltzoff, 2013), with some going so far as to suggest that mental state understanding (or at least its early precursors) evolved from (Blakemore & Decety, 2001; Frith & Frith, 1999) or is supported by (Marshall & Meltzoff, 2014; Woodward & Gerson, 2014) a neural system for detecting and representing actions.

A neural rhythm present in the EEG—known as the mu rhythm—has been postulated as a candidate neural mechanism facilitating developments from early emerging understanding of actions and intentions to more sophisticated representational understandings of complex mental states, including beliefs (e.g. Marshall & Meltzoff, 2014). In both adults and infants/children, mu rhythm reflects EEG oscillations (~8–13 Hz in adults, ~6–9 Hz in infants/children) recorded over and localized to sensorimotor cortex that suppress when adults, children, and infants either produce *or perceive* intentional actions (e.g., voluntary hand movement; see Fox et al., 2015, for meta-analysis; Hobson & Bishop, 2016, for large-scale study with adults). These findings suggest that mu suppression may be a reliable neural correlate of the representation of intentional action (e.g., Marshall & Meltzoff, 2011). It is these connections to both action production and action representation that lead researchers to postulate mu rhythm as a potential neural mechanism facilitating links between actions and the mental states that motivate them (Marshall & Meltzoff, 2014).

Recent evidence suggests that mu suppression is associated with the perception of meaningful intentional action in infants and with theory of mind developments in older children. For example, Filippi and colleagues (2016) examined mu rhythm as 7-month-old infants engaged in a goal-imitation paradigm. Infants viewed an actor reach for and grasp one of two objects and then had the opportunity to reach for and grasp either the same object previously chosen by the actor (goal action), or the unchosen object (non-goal action). Infants who showed strong mu suppression while watching the actor choose a toy were more likely to imitate the actor's goal action than were children who showed weak mu suppression. With preschool-age children, Bowman, Thorpe, Cannon, and Fox (2016) found that mu suppression was a critical factor in mediating the association between preschoolers' action production and their performance on standard theory of mind tasks. In each case, the findings suggest that mu suppression indexes processes that are critical for facilitating the links between mental states and action. Developing a detailed accounting of the role of mu suppression is an important avenue for understanding the online processes that are critical for imputing mental states to others.

Summary

There is now evidence that the functional development of the theory of mind network, including regions of the DMPFC and the rTPJ, are associated with transitions in preschoolers' theory of mind development. There is also evidence that the rTPJ, which becomes specialized to support reasoning about representational mental states, such as belief, has a longer developmental timeline than the DMPFC, which has a broader purpose in theory of mind reasoning and develops its specialization sooner. The importance of the DMPFC is underscored further by an association between dopaminergic functioning and theory of mind development. Emerging directions include understanding how the neural substrates that are responsive to the perception and production of intentional action also may be key contributors the emergence of theory of mind reasoning.

Experiential Bases of Theory of Mind Development

Although there are broad cross-cultural synchronies in theory of mind development that likely are shaped by specific neuromaturational factors, it now also is clear that variations in the rate and character of theory of mind development can be affected by children's sociocultural experiences. The role of experience is bluntly captured in training studies whereby children exposed to concentrated diets of theory of mind–relevant experience (one on one with an experimenter) show improved performance on tasks over a matter of minutes (e.g., Hale & Tager-Flusberg, 2003) or weeks (e.g., Amsterlaw & Wellman, 2006). However, in this section, we focus on research done within a sociocultural framework that views children as situated within a broad interconnected web of naturally occurring factors that can be more or less proximal to children's everyday experiences. (See, e.g., Carpendale & Lewis, 2004.) Despite their interconnected nature, we summarize the positive findings relevant to these factors independently. It is noteworthy that a recent meta-analysis testing effects to be described confirmed that the effects of most experiential factors are small (typically between $r = .15–.20$) but reliable across a number of studies (Devine & Hughes, 2016).

Parent–Child Conversation

Perhaps the most well-established finding in the literature on how experience affects theory of mind development is that parent talk about mental states is positively associated with preschool children's performance on false-belief tasks (e.g., Ruffman, Slade, & Crowe, 2002; see de Rosnay & Hughes, 2006, for a review). In these studies, parent–child dyads typically are observed in the laboratory or in their homes as they engage in an everyday activity, such as picture book reading or telling stories based on cards. These activities are thought to capture parents' "mind-mindedness"—or the extent to which they are focused on the appropriate (i.e., timely and accurate) discussion of mental states (Meins et al., 2002). These discussions can be about children's mental states, parents' mental states, or the mental states of the story characters they are reading about. Across a number of studies, parents' mental state talk, particularly their use of language relating to epistemic mental states, such as knowledge and beliefs, is associated positively with children's false-belief understanding (e.g., Adrian, Clemente, & Villanueva, 2007; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991; Racine, Carpendale, & Turnbull, 2006; Ruffman, Perner, & Parkin, 1999). Of these demonstrations, studies by Ruffman et al. (2002) and Adrían et al. (2007) were particularly notable because they employed longitudinal designs that enabled them to establish a causal role for parent mental state talk in the trajectory of false-belief understanding.

Siblings and Peer Interactions

Further evidence that is generally consistent with the view that conversational experience plays a critical role in theory of mind development comes from findings showing that the number of similar-age siblings preschoolers have is positively correlated with their false-belief understanding (e.g., Jenkins & Astington, 1996; Perner, Ruffman, & Leekam, 1994). In particular, it appears that the number of slightly older siblings is a critical factor (e.g., McAlister & Peterson, 2013). In each of these cases, it is thought that siblings provide preschoolers with regular opportunities to encounter mental states that are different from their own in ways that need to be understood in order for successful social interaction. Further to this point, Wang and Su (2009) recently found that false-belief understanding was accelerated in urban Chinese preschoolers (who at the time the research was done rarely had siblings) who were in daycare with children of varying ages rather than with same-age peers. These findings all support the general claim that a rich variety of conversational experience plays an important role in setting the timetable of theory of mind development.

Theory of Mind in Deaf Children

Perhaps the most dramatic demonstration of the role of verbal experience comes from work on individuals who are born deaf into hearing families. Children who are born deaf or who lose their hearing in the first year of life and are raised in in non-signing families do not experience the rich linguistic input that is immediately available to either deaf children raised in signing families or hearing children raised in speaking families (Peterson & Siegal, 1995). Several studies have demonstrated delays in the theory of mind performance of these deaf children. Compared to hearing children and deaf children born into signing families, deaf children born to non-signing families are substantially delayed on standard explicit false-belief tasks (Woolfe, Want, & Siegal, 2002), and on explicit understanding of other mental states, including desires and knowledge (Peterson, Wellman, & Liu, 2005). More recent work has suggested that these findings might be due to the fact that hearing parents of deaf children use fewer mental state terms than do hearing parents of hearing children (Moeller & Schick, 2006; Morgan et al., 2014). These findings show that just as typical variation in theory of mind–relevant conversational experience has small but reliable effects on preschoolers' false-belief development, the putatively extreme variation associated with deaf children born to hearing parents can result in striking delays.

Socioeconomic Status

In addition to the proximal factors that might affect the quantity and quality of children's theory of mind–relevant experiences directly, there is some evidence that socioeconomic status (SES) is positively associated with false-belief development as well. SES often is considered as an aggregate of a number of different variables, including parental education, income and occupation, marital status, and more. Although rarely an explicit focus of a study (but see Pears & Moses, 2003), SES measures often are collected by researchers interested in the effects of the experiential variables discussed earlier (e.g., Jenkins & Astington, 1996). In their meta-analysis, Devine and Hughes (2016) reported that, in fact, there is a highly reliable association between aggregate SES measures and false-belief understanding. Indeed, it is one of the strongest effects of experience reported. Thus, it appears that distal factors can have strong effects on the timetable and trajectory of theory of mind development.

Summary

Considerable evidence demonstrates the influence of children's social experience on their developing understanding of mental states. Through children's experiences with others—parents, siblings, same-age peers—an understanding of their own and others' minds is grown and sharpened. However, it is important to note that the meta-analytic data also demonstrates that the effects of these experiences are small. Except for the case of deaf children born to hearing parents, the correlations between social variables and children's false-belief understanding were statistically significant, but each accounted for just 2% and 4% of the variance in children's false-belief understanding performance. There are two important future directions for this literature. The first is that there may be much more work to do to better understand how experience affects theory of mind development. Although not many behavioral genetics studies have attempted to parse the extent to which theory of mind skills are heritable, one influential study suggests that a substantial proportion (44%) of the variance in 5-year-olds' theory of mind performance is due to nonshared theory of mind–specific experiences (Hughes et al., 2005). New estimates may suggest different targets, but specific work aimed at gaining a diverse set of reliable measures of that relevant experience may be critical to understanding its full effects and its theoretical importance. The second important direction is to recognize explicitly the interplay between these experiential factors and the cognitive architecture that is necessary to use that experience to promote meaningful conceptual change.

THEORY OF MIND IN THE REAL WORLD: A FEW EXAMPLES

From the outset, we noted that part of the appeal of theory of mind as a research topic is its importance in facilitating the smooth flow of everyday social interaction. The way in which we naturally make sense of others' behaviors is through understanding their intentions, which themselves are understood in terms of underlying beliefs and desires (e.g., Wellman, 1990). Because of this theoretical connection between theory of mind and social understanding, we would expect that facility with theory of mind reasoning might be predictive of social competence. As we consider this question, however, it is important to recognize that the links between theory of mind development and common metrics of social competence (e.g., peer acceptance, social network size, etc.) are likely to be weak and complex for several reasons. First, theory of mind skills may be particularly important for social problem solving but not for more affective aspects of social competence (e.g., empathy, homophily) that are also important for social success. Second, because of the changing nature of children's peer relationships, the importance of theory of mind may vary with age. Some work has attempted to connect theory of mind development with social competence, broadly construed. However, a tighter focus has been on detailing the connections between theory of mind development and those components of social understanding that especially require a mature understanding of mental states.

Theory of Mind and Social Competence

There is some evidence that individual differences in children's theory of mind are associated with social competence, broadly construed. For instance, one study found that the performance of 3- to 6-year-old children on false-belief tasks was positively correlated with teacher ratings of social skills but not with peer acceptance (Watson, Nixon, Wilson, & Cepage, 1999). When similar questions were investigated in another study that took a more fine-grained approach, false-belief understanding was associated with peer acceptance for older but not younger children (Slaughter, Dennis, & Pritchard, 2002). More recent findings suggest that the relationship may be a longitudinal one whereby early facility with theory of mind predicts later social success. For example, Devine, White, Ensor, and Hughes (2016) showed that theory of mind assessments given at age 6 were positively associated with teacher-rated social competence at age 10. (See also Caputi, Lecce, Pagnin, & Banerjee, 2012.) These findings suggest that at the broader level, theory of mind skills play some role in facilitating social competence over a long time scale but that meeting theory of mind milestones does not suddenly lead to increased social competence.

Deception

Deception is the intentional use of communicative acts to cause another to believe something that is not true. (See, e.g., Lee, 2013, for a recent review.) Deception and lying can happen for various purposes. For instance, some lies are self-serving—after causing damage to family property, a child might lie about involvement to avoid a potential punishment. Other lies might be undertaken to help another person feel good. For instance, upon receiving an undesirable item as a gift, one might lie about one's feelings in order to avoid offending the giver. Intuitively, successful deception appears to rely on false-belief understanding; after all, the point of deception is to cause someone else to hold a false belief about some true state of affairs (e.g., Newton, Reddy, & Bull, 2000). However, there are alternative explanations for children's early deception. For instance, it is possible that children who lie simply are hoping that disavowing responsibility for wrongdoing will help them avoid punishment, without understanding the causal role of false beliefs (e.g., Sodian, 1991).

One way in which researchers have moved away from this ambiguity in understanding the cognitive-motivational underpinnings of children's simple willingness to tell a lie is by investigating the more interesting question of whether developmental advances in false-belief understanding change the character of children's deception. That is, perhaps as children's theory of mind becomes more sophisticated, so too does their capacity for successful deception. Indeed, this seems to be the case. In a pioneering study, Talwar and Lee (2008) found that children's abilities to falsely deny that they had committed a minor transgression in a laboratory session (i.e., peeking at the contents of a container when they had been told not to) were positively related to their false-belief understandings. Furthermore, children's abilities to maintain a lie are associated with second-order theory of mind abilities; that is, children with stronger second-order false-belief skills were better able to conceal information that would be "telling" about their lie, such as telling facial and emotional expressions. (See Talwar, Gordon, & Lee, 2007.) Similar results have been demonstrated with lies to cover up transgressions and with lies that are meant to make others feel better (Broomfield, Robinson, & Robinson, 2002; Williams, Moore, Crossman, & Talwar, 2016). Thus, the data are clear that as children's theory of mind becomes more sophisticated, so too does their capacity for successful deception.

Moral Judgment

Theory of mind provides a way of explaining and predicting others' behaviors with respect to their most proximal causes (i.e., intentions, desires, and beliefs). Theory of mind considerations are also important more broadly for the meaningful evaluation of others' actions. One domain in which the evaluative importance of theory of mind is particularly apparent is in the case of making moral attributions. As philosophers have pointed out, others' actions are morally evaluated not simply in terms of their outcomes but also in terms of whether those outcomes were intended. (See, e.g., Kaplan, 2001; Mele & Sverdlik, 1996.) For instance, people intuitively assign blame and punishment more severely to people who intend to cause bad outcomes than they do to people who cause bad outcomes unintentionally (Leslie, Knobe, & Cohen, 2006). Indeed, the distinction between negative outcomes that are intended versus not intended is commonly included in most legal systems—when the negative outcome is the death of another individual, intentionality determines whether the crime is manslaughter or murder. Thus, our ability to make accurate inferences about others' intentions is a key feature of moral evaluation (Gray, Young, & Waytz, 2012; Knobe, 2005).

These intuitions about moral evaluation hinge on the ability to use an explicit representational theory of mind—that is, to think about how mental states are dissociable from observable actions. What we would predict, then, is that as children's theory of mind becomes more sophisticated, so too should their intuitions about moral responsibility. There is some evidence that around the time false-belief understandings are emerging, young children are more likely to judge negative outcomes that are brought about freely and intentionally as more immoral than ones that are brought about by force or accident (Chandler, Sokol, & Hallett, 2001; Josephs, Kushnir, Gräfenhain, & Rakoczy, 2016; Yuill & Perner, 1988). During the preschool years, children also come to believe that someone who attempted harmful behavior was morally wrong, even if the person was unsuccessful (Cushman, Sheketoff, Wharton, & Carey, 2013). More direct evidence on the question of the association between theory of mind development and morality comes from work by Killen and colleagues who presented children with what they called morally-relevant theory of mind tasks. In these tasks, children heard stories in which characters unknowingly and accidentally brought about a negative outcome, then children were asked to answer questions about the characters' intentions and beliefs. Children who correctly said that the story character was unaware that she had done something wrong also were less likely to view the transgression as morally wrong (Killen, Mulvey, Richardson, Jampol, & Woodward, 2011). Finally, there is evidence that theory of mind understandings measured in children 3.5 years old are positively associated with children's reasoning about moral acts 2 years later (Lane, Wellman, Olson, LaBounty, & Kerr, 2010).

Taken together, these findings clarify that children's developing understandings of the representational nature of mental states (as indexed by false-belief understanding) have important implications for their understanding of morality. Prior to the age of 4 years, children appear to believe that negative outcomes that are brought about either accidentally or intentionally are morally wrong. However, as they come to recognize that there can be a distinction between intentions and outcomes, they come to privilege intentions over outcomes in their assignment of moral value (Cushman et al., 2013).

Summary

There is evidence that theory of mind development is associated with social competence broadly construed and with specific social understandings that rely on the acquisition of particular theory of mind milestones—deception and moral judgement. Although we reviewed two specific social understandings, it should be noted that there is literature suggesting that theory of mind and its development might affect several other types of social understanding that have been explored, including playing strategic games (Sanfey, 2007), cooperation (Tomasello & Vaish, 2013), establishing trust (Mascaro & Sperber, 2009), and more. These bodies of work show that although theory of mind is a relatively narrowly construed aspect of conceptual development, its implications for children's abilities to navigate the social world are profound. This point becomes even clearer when considering the theory of mind difficulties experienced by those with particular developmental disorders and how those difficulties might affect their social cognitive functioning.

DEVELOPMENTAL DISORDERS, AUTISM, AND THEORY OF MIND

A host of developmental disorders are characterized by serious impairments in social functioning, including attention-deficit/hyperactivity disorder (ADHD), specific-language impairment, Down syndrome, and autism. Indeed, social impairment is often the most troubling symptom of these disorders for caregiving families. Given the hypothesized importance of theory of mind in social functioning, researchers have investigated whether difficulties in theory of mind might be a phenotypic characteristic of these developmental disorders. When investigating this question, however, there is an important distinction to be made with respect to characterizing the source of theory of mind difficulties. One possibility is that theory of mind understandings are affected in a domain-specific way, such that there are difficulties that exist for theory of mind reasoning independent of factors that also might affect conceptual understandings and problem solving in other domains, such as physical understandings or number. (See, e.g., Baron-Cohen, 1995.) An alternative possibility is that theory of mind difficulties exist because of primary impairments in domain-general factors that affect children's abilities either to develop or to use theory of mind in everyday situations. As was noted earlier, there is evidence that executive functioning is causally related to the theory of mind developments that occur over the preschool years. Accordingly, we might expect that any developmental disorder in which executive functioning or language development is impaired also would be associated with theory of mind difficulties and their sequelae.

Primary Deficits in Executive Function and Theory of Mind Delays

Although no one disorder has been studied in much detail, there is evidence that disorders that feature primary deficits in executive functioning skills also are associated with difficulties in false-belief reasoning; these disorders include Down syndrome (e.g., Cebula, Moore, & Wishart, 2010; Zelazo, Burack, Benedetto, & Frye, 1996), William syndrome (e.g., Mervis & John, 2010), fragile X syndrome (Grant, Apperly, & Oliver, 2007) and fetal alcohol spectrum disorder (Lindinger et al., 2016; Rasmussen et al., 2012). In the majority of these cases, the difference in false-belief performance between affected children and typically developing children is accounted for statistically by group differences in executive functioning. Moreover, in some cases, the association between executive functioning skills and false-belief performance is stronger in affected children relative to unaffected children (e.g., fetal alcohol spectrum disorder; see Rasmussen, Wyper, & Talwar, 2009).

A more mixed pattern of results is apparent for children with ADHD, which is another developmental disorder that is characterized by primary difficulties in executive function. Although some studies find that affected children have difficulties with theory of mind tasks relative to age-matched controls and that these differences can be accounted for by group differences in executive functioning (e.g., Mary et al., 2016), others have not (e.g., Charman, Carroll, & Sturge, 2001; Perner, Kain, & Barchfeld, 2002). The reasons for the mixed findings may have to do with the measures that are used; ADHD typically is not diagnosed until children enter school, and the common false-belief tasks that are used to assess theory of mind may not be appropriate for use with this age group. A recent study showed that although children with ADHD may not have difficulty with laboratory theory of mind tasks, they do have difficulties on tasks that require them to apply theory of mind understandings to real-world scenarios (Hutchins et al., 2016). Thus, evidence from ADHD may join evidence from other developmental disorders that are characterized by primary difficulties in executive functioning in showing that these conditions can be associated with secondary difficulties in theory of mind reasoning that may in turn affect aspects of social functioning.

Autism

Perhaps the developmental disorder that has received the most scrutiny in the present context is autism and autism spectrum disorders (ASDs; Baron-Cohen, 2000). Autism is a developmental disorder that has many features, but perhaps the most striking difficulty in autism concerns social impairments. This symptom profile led many researchers to suggest that theory of mind impairments may be a primary feature of ASD (Baron-Cohen, 1995) rather than a secondary feature that emerges because of other primary deficits such as executive function or language.

Even a cursory summary of the wide literature that has investigated theory of mind difficulties in autism is beyond the scope of this chapter. Briefly, the claim that individuals with ASD have a primary impairment in theory of mind reasoning comes from research that has taken a fine cuts approach. (See Frith & Happé, 1994). This approach is an experimental one in which researchers measure performance on theory of mind tasks—such as the false-belief task—along with performance on tasks that are formally and structurally similar to the false-belief task but involve nonmental content. For instance, Leslie and Thaiss (1992) measured the performance of individuals with ASD on a false-belief task and a "false photograph" task, which was thought to be identical to the false-belief task but involved nonmental content. Findings from these studies and others like them (e.g., Baron-Cohen et al., 1985) consistently find that ASD individuals perform well on tasks that involve reasoning about nonmental representations but poorly on tasks that involve reasoning about mental representations. (See, e.g., Baron-Cohen, 2000, for a review.) These results stand in contrast to typically developing individuals and those with non-ASD developmental disorders who tend to perform similarly on both kinds of tasks, in accordance with their mental age. However, recent research using a different nonmental representation task—the false sign task—calls into question the domain-specific theory of mind deficit in ASD. The false sign task is arguably more similar to the false-belief task because signs, like beliefs, are supposed to represent a current reality; whereas photographs are readily understood to have represented reality at a given point in the past. Children with ASD show difficulties in passing the false sign task (Iao & Leekam, 2014).

More generally, it should be noted that the theory of mind hypothesis of autism that proposes primary theory of mind deficits is controversial. A large literature shows that autism is associated with difficulties in executive functioning (e.g., Corbett, Constantine, Hendren, Rocke, & Ozonoff, 2009), so it remains possible that the theory of mind deficits in autism are secondary to those difficulties. Indeed, the association between executive functioning and false-belief performance is exceptionally high in ASD (Colvert, Custance, & Swettenham, 2002; Zelazo, Jacques, Burack, & Frye, 2002), which suggests that executive functioning may be the primary limiting factor on theory of mind reasoning in ASD. Others have suggested that theory of mind deficits in ASD may be secondary to a broader difficulty in forming central coherence or engaging in more holistic processing (Happé & Frith, 2006). Nonetheless, although it is still unclear whether theory of mind deficits are primary or secondary in autism, it is clear that individuals with ASD have substantial difficulty in reasoning about others' mental states.

Summary

There are developmental disorders in which theory of mind difficulties can be primary (such as autism) or secondary to deficits in cognitive skills that are important for acquiring or using theory of mind concepts (e.g., fetal alcohol spectrum disorder). Most important is that no matter what the source, difficulties in theory of mind reasoning may lead to broader difficulties with social functioning, which often is one of the most challenging aspects of caring for children with developmental delays. Taking a theory of mind perspective may be valuable insofar as it provides a specific target for intervention and training that could have broader implications for establishing more rewarding social interactions in these situations.

THEORETICAL FRAMEWORKS OF THEORY OF MIND DEVELOPMENT

The richness of the empirical work in theory of mind development also has been supported by a robust theoretical debate concerning the origins of theory of mind and the mechanisms by which theory of mind judgments become more reliable over time. Although a fair summary of the nuanced theoretical positions is well beyond the scope of this chapter, we attempt a brief summary of three general theoretical approaches here and highlight how each has inspired the empirical literature reviewed earlier in the chapter.

Theory Theory

The theoretical position that perhaps is most tightly connected with the theory of mind literature is the "theory" theory (Gopnik & Wellman, 1992, 2012; Perner, 1991). The premise of theory theory is that children begin life with a basic ability to impute general psychological causes to human behavior, but the precise content of the psychological causes is undifferentiated and/or immature. Over time and with experience, children are thought to notice instances in which their current understanding leads to mistaken or otherwise inept predictions of how people behave. These prediction errors serve as a catalyst for children to refine or revise their understandings about the nature of psychological causes, much in the same way that scientists refine or revise their theories about the causes of any natural phenomenon. In the end, these psychological causes are comprised of adult-like mental state concepts. Recently, computational processes (e.g., hierarchical Bayesian learning) have been offered as a means to characterize the ways in which the revision process might take place. Research summarized earlier in the chapter that has been spurred by the theory theory perspective includes (1) work that characterizes young children's (including infants') early understandings of mental states and maps out the qualitative differences between early and later understandings; (2) work that focuses on the ways in which children's diet of experience affects the timetable and trajectory of theory of mind development; and (3) work that characterizes the cognitive and neurobiological mechanisms that might be important for revising and refining mental state understandings based on experience in the world (e.g., executive functioning, MPFC, DA).

The general approach that theory theory takes to characterizing mechanisms of change is essentially continuous with longstanding theoretical perspectives in cognitive development, especially Piaget's processes of equilibration (e.g., assimilation and accommodation; Gopnik, 1996). Accordingly, the theory theory is susceptible to many of the critiques that applied to Piaget's model, including a lack of specificity of the processes by which change occurs and the characterization of the child as a highly rational agent who continually engages in sober reflection on mistaken ideas and their alternatives. (See, e.g., Faucher et al., 2002.) Nonetheless, the emphasis that the theory theory framework places on infants having some psychological understandings from early on that become elaborated over time and with experience has been a fruitful one for pursuing meaningful empirical work in the field.

Social Construction Views

Placing even more emphasis on the role of experience in theory of mind development are social construction views. Theorists here follow the lead of sociocultural theorists in arguing that ascribing mental states to others is a cultural practice and that the skillful application of those practices comes from cultural learning (e.g., Heyes & Frith, 2014). These intuitions have been formalized in theoretical accounts of theory of mind development. For example, Carpendale and Lewis (2004) made the case that theory of mind development takes place through social interaction. They argued that an understanding of the mind is constructed through triadic interactions among the child, another person, and the world. Nelson (2009) similarly argued that changes in children's understanding of mental states occur through pragmatic conversations in social interactions. Research carried out within this perspective concerns the effects of culture and experience on the timetable and trajectory of theory of mind development.

Although this perspective is an important one, it faces several challenges from the broader literature. First, it does not attempt to account for many of the known facts about the developmental trajectory of theory of mind understandings (e.g., why an understanding of desires comes before beliefs). Second, as noted by Devine and Hughes (2016) known experiential factors account for very little of the variance in preschoolers' theory of mind development. Nonetheless, this perspective is a valuable counterweight to the more child-focused processes that typically dominate theory and research in this area, and it reminds researchers to bear in mind the sociocultural context of socio-cognitive development.

Innateness and Modularity

A third prominent framework for studying theory of mind development is the innateness view. Proponents of the innate view suppose that the set of hypotheses for potential psychological causes cannot be constrained or specified on the basis of experience alone (German & Leslie, 2004). And thus, children must be born with mechanisms that take the perception of relevant social behaviors as input and generate a plausible set of psychological causes for those behaviors (Scholl & Leslie, 1999). Importantly, the character of those psychological causes is essentially adult-like from the beginning. Development occurs, then, as children become better at selecting the most appropriate psychological cause given the extant behavior. Research that has been spurred on by this perspective includes (1) work showing that infants exhibit false-belief understanding; (2) work showing that factors that affect the selection of appropriate psychological causes (e.g., executive functioning) affect both preschoolers' and adults' performance on theory of mind tasks (e.g., false belief); and (3) work showing that individuals with autism experience a specific deficit in reasoning about mental states, even when their reasoning in other domains is unaffected.

Critiques of the innateness view pertain largely to the fact that it does not easily account for particular findings in the literature, such as why young children proceed through a relatively stereotyped set of coherent understandings in a relatively circumscribed period later in development and why experience plays such a large role in shaping the timetable by which those understandings change (Wellman, 2014). Yet this approach has been extremely valuable for characterizing the remarkable sociocognitive competencies that infants show (irrespective of their ultimately correct interpretation) and the surprisingly specific ways in which theory of mind reasoning can be affected in cases of abnormal development.

CONCLUSION

Theory of mind is a research area that has seen sustained interest over the past 3 decades and continues to influence work on children's sociocognitive development. Arguably, we know more about theory of mind development during the preschool years than we know about any other comparable conceptual change. Here we have reviewed much of what is known about the timetable and trajectory of theory of mind development, the factors that affect those developments, and their broader consequences for social cognitive development. Although we have attempted to cover a lot of ground, we had to omit important work because of space considerations. Perhaps most notable among these are the rich and substantial literatures on connections between theory of mind and pretend play (see, e.g., Lillard et al., 2013, for a recent review; Taylor, Carlson, Maring, Gerow, & Charley, 2004) and between theory of mind and language (see, e.g., Astington & Baird, 2005; de Villiers, 2007). These literatures both have long empirical records and important theoretical implications, and we regret that we could not include substantive discussion of them here.

Although our field has acquired much important knowledge about theory of mind development, there still is much to learn about the developments that precede and succeed the milestones that occur in the preschool years and the mechanisms that support those changes. This research, particularly the research on mechanisms of change, still is in its early stages and holds substantial promise for understanding not just the particular developments that are important for theory of mind but conceptual development more broadly.

REFERENCES

Adrián, J. E., Clemente, R. A., & Villanueva, L. (2007). Mothers' use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: a longitudinal study. *Child Development*, *78*, 1052–1067.

Amsterlaw, J., & Wellman, H. M. (2006). Theories of mind in transition: a microgenetic study of the development of false-belief understanding. *Journal of Cognition and Development*, *7*, 139–172.

Apperly, I. A. (2010). *Mindreaders: The Cognitive Basis of Theory of Mind*. New York, NY: Psychology Press.

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: evidence from adult performance on a non-inferential theory of mind task. *Cognition*, *106*, 1093–1108.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970.

Apperly, I. A., Riggs, K. J., Simpson, A., Chiavarino, C., & Samson, D. (2006). Is belief reasoning automatic? *Psychological Science*, *17*, 841–844.

Astington, J. W., & Baird, J. A. (2005). *Why language Matters for Theory of Mind*. New York, NY: Oxford University Press.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*, 110–118.

Bamford, C., & Lagattuta, K. H. (2012). Looking on the bright side: children's knowledge about the benefits of positive versus negative thinking. *Child Development*, *83*, 667–682.

Banerjee, R., & Watling, D. (2005). Children's understanding of faux pas: associations with peer relations. *Hellenic Journal of Psychology*, *2*, 27–45.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S. (2000). Theory of mind and autism: a fifteen-year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience* (2nd ed., pp. 3–20). New York, NY: Oxford University Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37–46.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407–418.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: a study with normal adults, adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*, 241–251.

Bartsch, K., & Wellman, H. M. (1995). *Children Talk about the Mind*. New York, NY: Oxford University Press.

Benson, J. E., & Sabbagh, M. A. (2009). Theory of mind and executive functioning: a developmental neuropsychological approach. In P. D. Zelazo, E. Crone, & M. Chandler (Eds.), *Developmental Social Cognitive Neuroscience* (pp. 63–80). New York, NY: Oxford University Press.

Benson, J. E., Sabbagh, M. A., Carlson, S. M., & Zelazo, P. D. (2013). Individual differences in executive functioning predict preschoolers' improvement from theory of mind training. *Developmental Psychology*, *49*, 1615–1627.

Birch, S. A. J., Akmal, N., & Frampton, K. L. (2010). Two-year-olds are vigilant of others' non-verbal cues to credibility. *Developmental Science*, *13*, 363–369.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, *2*, 561–567.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false-belief task as a test of theory of mind. *Cognition*, *77*, B25–B31.

Bowman, L. C., Kovelman, I., Hu, X., & Wellman, H. M. (2015). Children's belief- and desire-reasoning in the temporoparietal junction: evidence for specialization from functional near-infrared spectroscopy. *Frontiers in Human Neuroscience*, *9*, 560. doi: 10.3389/fnhum.2015.00560.

Bowman, L. C., Liu, D., Meltzoff, A. N., & Wellman, H. M. (2012). Neural correlates of belief- and desire-reasoning in 7- and 8-year-old children: an event-related potential study. *Developmental Science*, *15*, 618–632.

Bowman, L. C., Thorpe, S. G., Cannon, E. N., & Fox, N. A. (2016). Action mechanisms for social cognition: behavioral and neural correlates of developing theory of mind. *Developmental Science*. doi: 10.1111/desc.12447

Boyes, M. C. & Chandler, M. (1992). Cognitive development, epistemic doubt, and identity formation in adolescence. *Journal of Youth and Adolescence*, *21*, 277–304.

Brandone, A. C., & Wellman, H. M. (2009). You can't always get what you want: infants understand failed goal-directed actions. *Psychological Science*, *20*, 85–91.

Broomfield, K. A., Robinson, E. J., & Robinson, W. P. (2002). Children's understanding about white lies. *British Journal of Developmental Psychology*, *20*, 47–65.

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., . . . Singh, S. (2005). Synchrony in the onset of mental-state reasoning: evidence from five cultures. *Psychological Science*, *16*, 378–384.

Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: the role of prosocial behavior. *Developmental Psychology*, *48*, 257–270.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*, 672–691.

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false-belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*, 1686–1706.

Carpendale, J. I., & Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *Behavioral and Brian Sciences*, *27*, 79–96.

Cebula, K. R., Moore, D. G., & Wishart, J. G. (2010). Social cognition in children with Down's syndrome: challenges to research and theory building. *Journal of Intellectual Disability Research*, *54*, 113–134.

Chandler, M. J., Sokol, B. W., & Hallett, D. (2001). Moral responsibility and the interpretive turn: children's changing conceptions of truth and rightness. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentions and Intentionality: Conceptions of Truth and Rightness* (pp. 345–365). Cambridge, MA: MIT Press.

Charman, T., Carroll, F., & Sturge, C. (2001). Theory of mind, executive function and social competence in boys with ADHD. *Emotional and Behavioral Difficulties*, *6*, 31–49.

Colvert, E., Custance, D., & Swettenham, J. (2002). Rule-based reasoning and theory of mind in autism: a commentary on the work of Zelazo, Jacques, Burack and Frye. *Infant and Child Development*, *11*, 197–200.

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research*, *166*, 210–222.

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, *34*, 28–38.

Couillard, N. L., & Woodward, A. L. (1999). Children's comprehension of deceptive points. *British Journal of Developmental Psychology*, *17*, 515–521.

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*, 6–21.

Deák, G. O., Flom, R., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, *36*, 511–523.

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*, 569–572.

De Rosnay, M., & Hughes, C. (2006). Conversation and theory of mind: do children talk their way to socio-cognitive understanding. *British Journal of Developmental Psychology*, *24*, 7–37.

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua*, *117*, 1858–1878.

Devine, R. T., & Hughes, C. (2014). Relations between false-belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*, 1777–1794.

Devine, R. T., & Hughes, C. (2016). Family correlates of false-belief understanding in early childhood: a meta-analysis. *Child Development*. doi: 10.1111/cdev.12682.

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*, 758–771.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*, 331–338.

Dunn, J., Brown, J. R., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Development*, *62*, 1352–1366.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*, 878–879.

Faucher, L., Mallon, R., Nazer, D., Nichols, S., Ruby, A., Stich, S., & Weinberg, J. (2002). The baby in the labcoat: why child development is an inadequate model for understanding the development of science. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The Cognitive Basis of Science* (pp. 335–362). Cambridge, UK: Cambridge University Press.

Filippi, C. A., Cannon, E. N., Fox, N. A., Thorpe, S. G., Ferrari, P. F., & Woodward, A. L. (2016). Motor system activation predicts goal imitation in 7-month-old infants. *Psychological Science*, *27*, 675–684.

Flavell, J., Flavell, E., & Green, F. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, *15*, 95–120.

Flynn, E. (2007). The role of inhibitory control in false-belief understanding. *Infant and Child Development*, *16*, 53–69.

Friedman, O., & Leslie, A. M. (2004). Mechanisms of belief-desire reasoning: inhibition and bias. *Psychological Science*, *15*, 547–552.

Frith, C. D., & Frith, U. (1999). Interacting minds: a biological basis. *Science*, *286*, 1692–1695.

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., . . . van Ijzendoorn, M. H. (2015). Assessing human mirror activity with EEG mu rhythm: a meta-analysis. *Psychological Bulletin*, *142*, 291–313.

Frith, U., & Happe, F. (1994). Autism: beyond "theory of mind." *Cognition*, *50*, 115–132.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483–527.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin*, *134*, 31–60.

German, T. P., & Leslie, A. M. (2004). No (social) construction without (meta-)representation: modular mechanisms as a *basis* for the capacity to acquire an understanding of mind. *Behavioral and Brain Sciences*, *27*, 106–107.

Gopnik, A. (1996). The post-Piaget era. *Psychological Science*, *7*, 221–225.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26–37.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, *7*, 145–171.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: causal models, Bayesian learning mechanisms and the theory theory. *Psychological Bulletin*, *138*, 1085–1108.

Grant, C. M., Apperly, I. A., & Oliver, C. (2007). Is theory of mind impaired in males with fragile X syndrome? *Journal of Abnormal Child Psychology*, *35*, 17–28.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101–124.

Gweon, H., Dodell-Feder, D., Bedny, M., & Saxe, R. (2012). Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child Development*, *83*, 1853–1868.

Gweon, H., & Saxe, R. (2013). Developmental cognitive neuroscience of theory of mind. In P. Rakic (Ed.), *Neural Circuit Development and Function in the Brain: Comprehensive Developmental Neuroscience* (Vol. 3, pp. 367–377). New York, NY: Elsevier.

Hala, S., Hug, S., & Henderson, A. (2003). Executive function and false-belief understanding in preschool children: two tasks are harder than one. *Journal of Cognition and Development*, *4*, 275–298.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on theory of mind: a training study. *Developmental Science*, *6*, 346–359.

Hamlin, J. K., Hallanan, E. V., & Woodward, A. L. (2008). Do as I do: 7-month-old infants selectively reproduce others' goals. *Developmental Science*, *11*, 487–494.

Happé, F., & Frith, U. (2006). The weak central coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*, 5–25.

Harris, P., Donnelly, K., Guz, G., & Pitt-Watson, R. (1986). Children's understanding of the distinction between real and apparent emotion. *Child Development*, *57*(4), 895–909. doi:10.2307/1130366

Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science*, *17*, 647–659.

Heyes, C., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*, 1243091.

Hobson, H. M., & Bishop, D. V. M. (2016). Mu suppression—A good measure of the human mirror neuron system? *Cortex*, *82*, 290–310.

Hughes, C. (1998). Executive functioning in preschoolers: links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*, 233–253.

Hughes, C., Jaffee, S. R., Happé, F., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Origins of individual differences in theory of mind: from nature to nurture? *Child Development*, *76*, 356–370.

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*, 590–619.

Hunnius, S., & Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philosophical Transactions of the Royal Society B*, *369*, 20130490. doi: 10.1098/rstb.2013.0490.

Hutchins, T. L., Prelock, P. A., Morris, H., Benner, J., LaVigne, T., & Hoza, B. (2016). Explicit vs. applied theory of mind competence: a comparison of typically developing males, males with ASD, and males with ADHD. *Research in Autism Spectrum Disorders*, *21*, 94–108.

Iao, L. S., & Leekam, S. R. (2014). Nonspecificity and theory of mind: new evidence from a nonverbal false-sign task and children with autism spectrum disorders. *Journal of Experimental Child Psychology*, *122*, 1–20.

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, *32*, 70–78.

Josephs, M, Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology*, *141*, 247–255.

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications*, *2*, 341. doi: 10.1038/ncomms1342

Kaplan, L. V. (2001). Intentional agency, responsibility, and justice. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.) *Intentions and Intentionality* (pp. 367–379). Cambridge, MA: MIT Press.

Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: morally-relevant theory of mind. *Cognition*, *119*, 197–215.

Knobe, J. (2005). Theory of mind and moral cognition: exploring the connections. *Trends in Cognitive Sciences*, *9*, 357–359.

Koster-Hale, J., & Saxe, R. (2013). Functional neuroimaging of theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & M. Lombardo (Eds.), *Understanding Other Minds* (3rd ed., pp. 132–163). New York, NY: Oxford University Press.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science*, *330*, 1830–1834.

Lackner, C. L., Bowman, L. C., & Sabbagh, M. A. (2010). Dopaminergic functioning and preschoolers' theory of mind. *Neuropsychologia*, *48*, 1767–1774.

Lackner, C. L., Sabbagh, M. A., Hallinan, E. V., Liu, X., & Holden, J. J. E. (2012). Dopamine receptor D4 gene variation predicts preschoolers' developing theory of mind. *Developmental Science*, *15*, 272–280.

Lagattuta, K. H., Kramer, H. J., Kennedy, K., Hjortsvang, K., Goldfarb, D., & Tashjian, S. (2015). Beyond Sally's missing marble: further development in children's understanding of mind and emotion in middle childhood. *Advances in Child Development and Behavior*, *48*, 185–217.

Lagattuta, K. H., Sayfan, L., & Blattman, A. J. (2010). Forgetting common ground: six- to seven-year-olds have an over-interpretive theory of mind. *Developmental Psychology*, *46*, 1417–1432.

Lagattuta, K. H., Sayfan, L., & Harvey, C. (2014). Beliefs about thought probability: evidence for persistent errors in mindreading and links to executive control. *Child Development*, *85*, 659–674.

Lane, J. D., Wellman, H. M., Olson, S. L., LaBounty, J., & Kerr, D. C. (2010). Theory of mind and emotion understanding predict moral development in early childhood. *British Journal of Developmental Psychology*, *28*, 871–889.

Lee, K. (2013). Little liars: development of verbal deception in children. *Child Development Perspectives*, *7*, 91–96.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind." *Trends in Cognitive Sciences*, *8*, 528–533.

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect. *Psychological Science*, *17*, 421–427.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition*, *43*, 225–251.

Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology*, *4*, 231–244.

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: a review of the evidence. *Psychological Bulletin*, *139*, 1–34.

Lindinger, N. M., Malcolm-Smith, S., Dodge, N. C., Molteno, C. D., Thomas, K. G., Meintjes, E. M., . . . Jacobson, S. W. (2016). Theory of mind in children with fetal alcohol spectrum disorders. *Alcoholism, Clinical and Experimental Research*, *40*, 367–376.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, *108*, 732–739.

Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2009). Neural correlates of theory of mind reasoning in adults and children. *Child Development*, *80*, 318–326.

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*, 523–531.

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: a primer on the two-systems account. *Child Development Perspectives*, *10*, 184–189.

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*, 305–311.

Marshall, P. J., & Meltzoff, A. N. (2011). Neural mirroring systems: exploring the EEG mu rhythm in human infancy. *Developmental Cognitive Neuroscience*, *1*, 110–123.

Marshall, P. J., & Meltzoff, A. N. (2014). Neural mirroring mechanisms and imitation in human infants. *Philosophical Transactions of the Royal Society B*, *369*, 20130181. doi: 10.1098/rstb.2013.0181.

Mary, A., Slama, H., Mousty, P., Massat, I., Capiau, T., Drabs, V., & Peigneux, P. (2016). Executive and attentional contributions to theory of mind deficit in attention deficit/hyperactivity disorder. *Child Neuropsychology*, *22*, 345–365.

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, *112*, 367–380.

McAlister, A. R., & Peterson, C. C. (2013). Siblings, theory of mind, and executive functioning in children aged 3–6 years old: new longitudinal evidence. *Child Development*, *84*, 1442–1458.

Meinhardt, J., Sodian, B., Thoermer, C. Döhnel, K., & Sommer, M. (2011). True- and false-belief reasoning in children and adults: an event-related potential study of theory of mind. *Developmental Cognitive Neuroscience*, *1*, 67–76.

Meins, E., Fernyhough, C., Wainwright, R., Das Gupta, M., Fradley, E., & Tuckey, M. (2002). Maternal mind-mindedness and attachment security as predictors of theory of mind understanding. *Child Development*, *73*, 1715–1726.

Mele, A., & Sverdlik, S. (1996). Intention, intentional action, and moral responsibility. *Philosophical Studies*, *82*, 265–287.

Meltzoff, A. N. (2013). Origins of social cognition: bidirectional self-other mapping and the "like-me" hypothesis. In M. Banaji & S. Gelman (Eds.), *Navigating the Social World: What Infants, Children and Other Species Can Teach Us* (pp. 139–144). New York, NY: Oxford University Press.

Mervis, C. B., & John, A. E. (2010). Cognitive and behavioral characteristics of children with Williams syndrome: implications for intervention approaches. *American Journal of Medical Genetics*, *154C*, 229–248.

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin*, *135*, 749–773.

Moeller, M. P., & Schick, B. (2006). Relations between maternal input and theory of mind understanding in deaf children. *Child Development*, *77*, 751–766.

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*, F1–F9.

Moore, C. (1999). Gaze following and the control of attention. In P. Rochat (Ed.), *Early Social Cognition: Understanding Others in the First Month of Life* (pp. 241–256). Mahwah, NJ: Erlbaum.

Morgan, G., Meristo, M., Mann, W., Hjelmquist, E., Surian, L., & Siegal, M. (2014). Mental state language and quality of conversational experience in deaf and hearing children. *Cognitive Development*, *29*, 41–49.

Mosconi, M. W., Mack, P. B., McCarthy, G., & Pelphrey, K. A. (2005). Taking an "intentional stance" on eye-gaze shifts: a functional neuroimaging study of social perception in children. *NeuroImage*, *27*, 247–252.

Moses, L. J. (2001). Executive accounts of theory of mind development. *Child Development*, *72*, 688–690.

Moses, L. J., & Sabbagh, M. A. (2007). Interactions between domain-specific and domain-general processing in the development of children's theory of mind. In M. J. Roberts (Ed.), *Integrating the Mind: Domain-General versus Domain-specific Processes in Higher Cognition* (pp. 275–291). New York, NY: Psychology Press.

Nelson, K. (2009). Narrative practices and folk psychology: a perspective from developmental psychology. *Journal of Consciousness Studies*, *16*, 69–93.

Newton, P., Reddy, V., & Bull, R. (2000). Children's everyday deception and performance on false-belief tasks. *British Journal of Developmental Psychology*, *18*, 297–317.

Oh, S., & Lewis, C. (2008). Korean preschoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development*, *79*, 80–99.

Ohnishi, T., Moriguchi, Y., Matsuda, H., Mori, T., Hirakata, M., Imabayashi, E., . . . Uno, A. (2004). The neural network for the mirror system and mentalizing in normally developed children: An fMRI study. *Neuroreport*, *15*(9), 1483–1487.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258.

Pears, K. C., & Moses, L. J. (2003). Demographics, parenting and theory of mind in preschool children. *Social Development*, *12*, 1–20.

Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Perner, J., Kain, W., & Barchfeld, P. (2002). Executive control and higher-order theory of mind in children at risk of ADHD. *Infant and Child Development*, *11*, 141–158.

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: you catch it from your sibs. *Child Development*, *65*, 1228–1238.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that . . . ": attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471.

Peterson, C. C., & Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry*, *36*, 459–474.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, *76*, 502–517.

Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D. (2009). Neural correlates of direct and reflected self-appraisals in adolescents and adults: when social perspective-taking informs self-perception. *Child Development*, *80*, 1016–1038.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*, 53–78.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not level-1 visual perspective calculation: evidence from a dual-task study with adults. *Cognition*, *117*, 230–236.

Racine, T. P., Carpendale, J. I. M., & Turnbull, W. (2006). Cross-sectional and longitudinal relations between mother-child talk about conflict and children's social understanding. *British Journal of Psychology*, *97*, 521–536.

Rakoczy, H, Fizke, E., Bergfeld, D., & Schwarz, I. (2015). Explicit theory of mind is even more unified than previously assumed: belief ascription and understanding aspectuality emerge together in development. *Child Development*, *86*, 486–502.

Rasmussen, C., Tamana, S., Baugh, L., Andrew, G., Tough, S., & Zwaigenbaum, L. (2012). Neuropsychological impariments on the NEPSY-II among children with FASD. *Child Neuropsychology*, *19*, 337–349.

Rasmussen, C., Wyper, K., & Talwar, V. (2009). The relation between theory of mind and executive functions in children with fetal alcohol spectrum disorders. *Canadian Journal of Clinical Pharmacology*, *16*, e370–e380.

Reddy, V. (1991). Playing with others' expectations: teasing and mucking about in the first year. In A. Whiten (Ed.), *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading* (pp. 143–158). Oxford, UK: Blackwell.

Reddy, V. (2007). Getting back to the rough ground: deception and "social living." *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 621–637.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*, 12–21.

Robson, S. J., & Kuhlmeier, V. A. (2016). Infants' understanding of object-directed action: an interdisciplinary synthesis. *Frontiers in Psychology*. doi:10.3389/fpsyg.2016.00111

Ross, H. S., Recchia, H. E., & Carpendale, J. I. M. (2005). Making sense of divergent interpretations of conflict and developing an interpretive theory of mind. *Journal of Cognition and Development*, *6*, 571–592.

Ruffman, T. (2014). To belief or not belief: children's theory of mind. *Developmental Review*, *34*, 265–293.

Ruffman, T., Perner, J., & Parkin, L. (1999). How parenting style affects false-belief understanding. *Social Development*, *8*, 395–411.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*, 734–751.

Russell, J. A. (1996). *Agency: Its Role in Mental Development*. Mahwah, NJ: Erlbaum.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172.

Sabbagh, M. A. (2004). Recognizing and reasoning about mental states: Understanding orbitofrontal contributions to theory of mind and autism. *Brain and Cognition*, *55*, 209–219.

Sabbagh, M. A. (2013). Brain electrophysiological studies of theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & M. Lombardo (Eds.) *Understanding Other Minds* (3rd ed., pp. 119–131). New York, NY: Oxford University Press.

Sabbagh, M. A. (2016). Dopamine and theory of mind development. In J. Sommerville & J. Decety (Eds.), *Social Cognition: Frontiers in Developmental Science* (pp. 187–209). New York, NY: Psychology Press.

Sabbagh, M. A., Benson, J. E., & Kuhlmeier, V. A. (2013). False-belief understanding in infants and preschoolers. In M. Legerstee, D. Haley, & M. Bornstein (Eds.), *The Infant Mind: Origins of the Social Brain* (pp. 301–323). New York, NY: Guilford Press.

Sabbagh, M. A., Bowman, L. C., Evraire, L., & Ito, J. M. B. (2009). Neurodevelopmental bases of preschoolers' theory of mind development. *Child Development*, *80*, 1147–1162.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: a comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*, 74–81.

Sanfey, A. G. (2007). Social decision-making. Insights from game theory and neuroscience. *Science*, *318*, 598–602.

Saxe, R. R., Whitfield-Gabrieli, S., Scholz, J., & Pelphrey, K. A. (2009). Brain regions for perceiving and reasoning about other people in school-aged children. *Child Development*, *80*, 1197–1209.

Sayfan, L., & Lagattuta, K. H. (2008). Grown-ups are not afraid of scary stuff, but kids are: young children's and adults' reasoning about children's, infants' and adults' fears. *Child Development*, *79*, 821–835.

Schneider, D., & Low, J. (2016). Efficient versus flexible mentalizing in complex social settings: exploring signature limits. *British Journal of Psychology*, *107*, 26–29.

Scholl, B. J., & Leslie, A. M. (1999). Modularity, development, and "theory of mind." *Mind and Language*, *14*, 131–153.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, *80*, 1172–1196.

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*, 1239–1247.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*, 617–627.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: a systematic investigation of the first reference to mental state. *Cognition*, *14*, 301–321.

Shimizu, Y., & Johnson, S. C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science*, *7*, 425–430.

Siegel, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition*, *38*, 1–12.

Slaughter, V., Dennis, M. J., & Pritchard, M. (2002). Theory of mind and peer acceptance in preschool children. *British Journal of Developmental Psychology*, *20*, 545–564.

Sodian, B. (1991). The development of deception in young children. *British Journal of Developmental Psychology*, *9*, 173–188.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*, B1–B11.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*, 587–592.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*, 580–586.

Surian, L., & Geraci, A. (2012). Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *British Journal of Developmental Psychology*, *30*, 30–44.

Talwar, V., Gordon, H. M., & Lee, K. (2007). Lying in the elementary school years: verbal deception and its relation to second-order belief understanding. *Developmental Psychology*, *43*, 804–810.

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, *79*, 866–881.

Tardif, T., Wellman, H. M., & Cheung, K. M. (2004). False-belief understanding in Cantonese-speaking children. *Journal of Child Language*, *31*, 779–800.

Taylor, M., Carlson, S. M., Maring, B. L., Gerow, L., & Charley, C. M. (2004). The characteristics and correlates of fantasy in school-age children: imaginary companions, impersonation, and social understanding. *Developmental Psychology*, *40*, 1173–1187.

Thoermer, C., Woodward, A., Sodian, B., Perst, H., & Kristen, S. (2013). To get the grasp: seven-month-olds encode and selectively reproduce goal-directed grasping. *Journal of Experimental Child Psychology*, *116*, 499–509.

Tomasello, M. (2014). *A Natural History of Human Thinking*. Cambridge, MA: Harvard University Press.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology*, *39*, 906–912.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, *64*, 231–255.

Wang, Y., & Su, Y. (2009). False-belief understanding: children catch it from classmates of different ages. *International Journal of Behavioral Development*, *33*, 331–336.

Watson, A. C., Nixon, C. L., Wilson, A., & Cepage, L. (1999). Social interaction skills and theory of mind in young children. *Developmental Psychology*, *35*, 386–391.

Wellman, H. M. (1990). *The Child's Theory of Mind*. Cambridge, MA: MIT Press.

Wellman, H. M. (2002). Understanding the psychological world: developing a theory of mind. In. U. Goswami (Ed.), *Blackwell Handbook of Childhood Cognitive Development* (pp. 167–187). Malden, MA: Blackwell.

Wellman, H. M. (2014). *Making Minds: How Theory of Mind Develops*. New York, NY: Oxford University Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: the truth about false belief. *Child Development*, *72*, 655–684.

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling theory-of-mind understandings in Chinese children. *Psychological Science*, *17*, 1075–1081.

Wellman, H. M., & Liu, D. (2004). Scaling theory-of-mind tasks. *Child Development*, *75*, 523–541.

Westra, E. (2016). Pragmatic development and the false-belief task. *Review of Philosophy and Psychology*. doi:10.1007/s13164-016-0320-5

Williams, S., Moore, K., Crossman, A. M., & Talwar, V. (2016). The role of executive functions and theory of mind in children's prosocial lie telling. *Journal of Experimental Child Psychology*, *141*, 256–266.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: theory of mind in deaf children. *Child Development*, *73*, 768–778.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Woodward, A. L., & Gerson, S. A. (2014). Mirroring and the development of action understanding. *Philosophical Transactions of the Royal Society B*, *369*, 20130181. doi:10.1098/rtsb.2013.0181

Yuill, N., & Perner, J. (1988). Intentionality and knowledge in children's judgments of actors' responsibility and recipients' emotional reactions. *Developmental Psychology*, *24*, 358–365.

Zelazo, P. D., Burack, J. A., Benedetto, E., & Frye, D. (1996). Theory of mind and rule use in individuals with Down syndrome: a test of the uniqueness and specificity claims. *Journal of Child Psychology and Psychiatry*, *37*, 479–484.

Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (2nd ed., pp. 553–574). Cambridge, MA: MIT Press.

Zelazo, P. D., Jacques, S., Burack, J. A., & Frye, D. (2002). The relation between theory of mind and rule use: evidence from persons with autism-spectrum disorders. *Infant and Child Development*, *11*, 171–195.

CHAPTER 9

Emotion Development from an Experimental and Individual Differences Lens

KORALY PÉREZ-EDGAR AND PAUL HASTINGS

INTRODUCTION

Developmental psychology, like the broader science of psychology, sets out to capture a phenomenon of interest in a moment in time, noting both the central tendency (the normative centroid) and variation surrounding the center. From this point, however, developmental psychology sets itself a bit apart in its added motivation to build from a single snapshot to capture both change over time and the mechanisms that fuel change. This scientific worldview is applied across a wide spectrum of phenomena, including behavior, cognition, and affect, and across multiple layers of analysis. At times, the field has struggled with how to best empirically translate this shared worldview. Thus, there have been, and continue to be, deep discussions on the role of metatheory and methods, debates as to whether we are a science of the laboratory or the field, and arguments characterizing the role of developmental psychologists as quiet anthropological observers or the active tinkerers of the chemistry lab.

It is in this context that the current chapter examines the methodological and theoretical considerations at play in the study of emotion development. A number of crucial reviews (Denham, 1998; Ekman, 1999; Harris, 1989; M. Lewis, 2010) examine the nature and teleology of emotion, carefully charting the rise and transformation of emotion(s) across development and contexts. This is not one of them. Rather, this chapter focuses on the approaches developmental psychologists have taken in examining emotion, both historically and today. In doing so, we document an ongoing struggle to place experimental work within our subfield of affective developmental science. Here we use the specific term "affective developmental science" as it reflects the (current) focus on capturing and explaining variance in emotion across the life span using a broad spectrum of empirical tools. Not surprisingly, we argue that experimental work has an important role in shaping our understanding of socioemotional variation previously observed in naturalistic settings while also providing the vital clues needed to inform future observational work. Although this may not appear to be a controversial statement, the field often has been suspicious of this approach, remaining on guard, lest developmental psychology fall into the trap of "deifying the manipulative experimental method" (McCall, 1977, p. 336).

In discussing the role of experimental methods in affective developmental science, we also note a research tradition that has received relatively less attention: the identification and systematic study of individual differences. To some extent, both naturalistic and experimental researchers share a faint disdain for individual differences. The naturalistic approach, in its purest incarnation, wishes to capture near-universal developmental pathways (Wohlwill, 1973), while the experimental approach wishes to impose strict control such that the only variation present builds directly from the clever manipulations imposed by the researcher (Reese & Lipsitt, 1970). However, we suggest that individual differences, systematically studied rather than systematically removed, play an important role in delineating mechanisms of socioemotional development, help define what we mean by "normative," reveal the ways in which individual traits come together with environmental context to shape trajectories, and focus our resources on the strongest targets for identification, prevention, and intervention.

Thus, this chapter looks to see how experimental methods, coupled with an eye to capturing individual differences, may help expand our understanding of emotion and socioemotional development. We base this discussion on the premise that emotions are biologically prepared adaptive processes (P. M. Cole, Martin, & Dennis, 2004; Dennis, Buss, & Hastings, 2012; Hastings, Miller, Kahle, & Zahn-Waxler, 2014). From this perspective, emotions are integral to generating a goal, maintaining progress toward the goal, and assessing the impact of attaining, or failing to attain, the goal. Emotions are pervasive across development. They draw from, and are reflected in, neurobiological, perceptual, cognitive, and behavioral systems (Thompson, 2011). It is the multidimensional nature of emotion that most clearly calls for a multidimensional approach that incorporates both descriptive and experimental work and both a normative and an individual differences lens.

In this chapter we also lean heavily on the temperament literature, using the construct as a model system to illustrate the promise (and limitations) of the experimental and individual differences approach to studying emotion development. In part, this reflects the authors' biases in the work they have carried out to date. However, we would like to believe that this emphasis also reflects a real (if perhaps unintended) tradition within the study of temperament of drawing from multiple research streams to study a core question centered on the early and evolving mechanisms that shape developmental pathways from infancy through senescence.

Specifically, we draw on the literature examining behavioral inhibition. Kagan and colleagues (Garcia Coll, Kagan, & Reznick, 1984; Kagan, 1994, 2012) first described the temperamental trait of behavioral inhibition in children. As infants, behaviorally inhibited children display signs of fear and wariness in response to unfamiliar stimuli (Schmidt et al., 1997), and this trait is marked by heightened vigilance, motor quieting, and withdrawal from novelty (Garcia Coll et al., 1984; Kagan, Reznick, & Snidman, 1987). By elementary school, many behaviorally inhibited children fear social circumstances, displaying poorly regulated social behavior and social reticence (Coplan, Rubin, Fox, Calkins, & Stewart, 1994; N. A. Fox et al., 1995). This difficulty with social interaction, in turn, increases the likelihood of peer rejection, low self-esteem, poor social competence, and even academic difficulties (Hastings, Kahle, & Nuselovici, 2014; Rubin, Chen, & Hymel, 1993; Schmidt, Fox, Schulkin, & Gold, 1999). Longitudinal studies of behavioral inhibition, and the broader construct of temperamental shyness, have found a marked increased risk for anxiety, particularly social anxiety, by midadolescence (Chronis-Tuscano et al., 2009; Clauss & Blackford, 2012; Kagan, Snidman, McManis, & Woodward, 2001).

Of particular interest to this chapter, the construct of behavioral inhibition first emerged from careful observation of children both in their natural environment and in the laboratory. Experimental studies then probed the mechanisms that led to observed profiles of behavioral inhibition and underscored the link to socioemotional outcomes (Kagan, 2012). Temperament, as a construct, also argues for the importance of individual differences in children's experience of and response to their social worlds. Thus, this area of research will help illustrate many of the core arguments touched on in this chapter.

HISTORICAL TRENDS SHAPING THE STUDY OF EMOTION DEVELOPMENT

What Does It Mean to Be a "Real" Science?

The subfield of affective developmental science did not emerge, in either name or form, from a historical vacuum. Indeed, the name itself reflects current trends regarding how to best approach centuries-old questions regarding development. Overton (2006) argued that every field has meta-theories that shape trends in research across subdisciplines. Meta-theories provide the concepts and contexts from which specific theories and methods can emerge. That is, they create the narrative that shape the questions asked by the field. Overton suggested that the current meta-theory that defines the scientific method is rooted in "observation, causation, and induction-deduction" (p. 19). Within that, more localized historical forces can shape specific theories central to the field. Meta-theories then lead to meta-methods—acceptable strategies for answering questions of interest. For example, the dominance of behaviorism in the early 20th century reflected the shared understanding that the environment was the primary force shaping developmental trajectories (Watson, 1926). This view was then supplemented (and perhaps supplanted) by biologically based mechanisms and cognitive processes (Bjorklund, 1997) that brought the individual child into the developmental process.

Researchers tend to accept and embrace tools that are most adequate to explore the world described by the meta-theory. The use of acceptable methods—rooted in acceptable theories—helps us both judge the quality of any one study and place the field within the larger community of science. Today's affective developmental science reflects these historical trends.

Cronbach (1957) painted a vivid picture of the various forces coming together to create a modern science of psychology. In his presidential address to the American Psychological Association, he noted that the field was split into two competing factions. Experimental psychology was a tight little island, with clear borders and strict rules of admission. "Correlational" psychology, in contrast, was the Holy Roman Empire—large, sprawling, and perhaps ultimately ungovernable. In his estimation, the well-guarded island was "much the more coherent of our two disciplines," This was all for the better since "it is these methods which qualify us as scientists, rather than philosophers or artists" (p. 671). This fit with psychology's striving to display its scientific bona fides, so that it would be deemed worthy companions to the hard(er) sciences of physics, biology, and chemistry.

The historical trend was clear. Any discipline, or subdiscipline, of psychology would need to fit these strictures in order to be allowed into the fold. For example, Cronbach (1957) singled out Harlow as creating a truly experimental psychology of development. In Harlow's work, one could see tight control of the organism's environment, the clear definition of the outcome (dependent) variable of interest, and the systematic manipulation of processes (the independent variable) contributing to a rich mechanistic view of socioemotional development (Harlow, 1958).

The field of "child psychology" was critiqued for having few experimental controls and techniques, being frequently atheoretical, and trying to capture complex variables that were difficult to define or control (Wohlwill, 1973). The men and women interested in the developing child were very much aware of these pervasive critiques and acted accordingly, moving away from naturalistic observational methods. In response, child psychology gave way to a new "developmental psychology" that looked to experimental science as its brethren and empirical ideal.

These trends were evident as early as the 1930s, when the historical antecedents to today's developmental psychology worked to place themselves under the umbrella of experimental psychology, distinguishing themselves from prior impressionistic studies and popular "baby diaries" (Wallace, Franklin, & Keegan, 1994). In doing so, researchers often took on the traditional topics of experimental psychology, simply replacing adults with children as the organism of study. The long line of studies in children's learning and learning theory (Reese & Lipsitt, 1970) fit this mold. As McCall (1977) noted, researchers were working to "infuse developmental psychology with scientific respectability" (p. 333).

However, by the 1970s, many leaders of the field fretted that a "substantial science of naturalistic developmental processes" was lacking (McCall, 1977, p. 333) and that developmental psychology had suffered in "the invasion of the experimentalists" (Wohlwill, 1973, p. 8). Wohlwill (1973) argued that if developmental psychology did not reembrace a focus on development, the field would lose its place as a distinct contributor to psychology. Rather, it would devolve into a paler branch of general psychology defined simply by the age of the participants. Wohlwill's concern resulted in a call to arms, exhorting the new generation of developmental psychologists to embrace the longitudinal study. Researchers were to invest the time and effort to create sustained and careful description of children's natural trajectories. This naturalistic focus was to distinguish developmental psychology, and developmental psychologists, from the mechanistic tinkerers of the other subdisciplines.

In the last four decades, the field has embraced the ideal so wholeheartedly that today many discussion sections include the mea culpa limitation of being cross-sectional, experimental, or both. The experimentalists took a strategic retreat to focus on infancy—and, in particular, cognitive development in infancy (Aslin & Fiser, 2005). Work focused on socioemotional development was "to remain at an essentially descriptive level" (Wohlwill, 1973, p. 14). The methodological exclusivity was bidirectional. For example, Reese and Lipsitt (1970) wrote a comprehensive 700+-page book reviewing experimental child psychology. In it, they devoted 45 pages to set transfer in learning and 10 pages to all of emotional development.

The distinction in approach, and the animosity among proponents, often was attributed to differences in training and scientific background. Boring (1929) went so far as to suggest that the divide was rooted in personality traits clustering in the two camps. Whatever trait-level differences were evident, they were then magnified by the specific training histories within the camps. Specific developmental curricula in graduate training programs emerged in the 1970's and 1980's, narrowing the range of variables and processes related to child development examined by any one group of researchers. As a result, each camp thought the other vaguely scientifically suspect. This suspicion echoed Catell's (1898) earlier concern that, after training as experimental psychologists, "regard for the body of nature becomes that of the anatomist, rather than that of the lover" (p. 152).

A shared common interest in variation did manage to bridge the methodological divide. Although the experimentalist was thought to focus solely on the variation he or she created, the correlational psychologist was interested in already existing variation to be found in nature. Ironically, the shared focus on variation also led to a shared antipathy toward individual differences.

Within the correlational camp, the focus was on variation across a central tendency that is presumed to be somewhat linear and driven by a single, shared, causal mechanism. These developmental functions capture the amount or frequency of a behavior across age for an individual or group of individuals. Indeed, McCall (1977) argued for a descriptive science of developmental function rather than prediction. When individual differences were invoked, the focus was on relative rank ordering for an individual on a given attribute relative to rank ordering on the same trait at a different time. In this view, "correlation" implies stability of individual differences. This approach was not meant to embrace a diversity of mechanism or pathway.

Experimentalists did not have much more use for individual differences in their work, as they were seen as the enemy of experimental control. Individual differences were the marker of sloppy methodology and agents obscuring functional mechanisms. Indeed, Watson (1926) argued that if you can experimentally make and unmake individual differences at will, then by definition the differences have little scientific importance. The experimental ideal was to identify and manipulate a strong "treatment" that could reliably and universally bring about a predicted pattern of behavior. This point of view failed to highlight that the capacity to show variation, either at the level of a specific trait or in a class of individuals, may in and of itself provide needed insight into a construct of interest and openness to developmental change.

These historical trends, although perhaps not as starkly etched, still are evident in our modern science. As such, they shape both the questions asked in developmental affective science and how we go about answering these questions.

What Are the Core Methodologies in Developmental Science?

Developmental science is focused on determining causality, capturing a profile or relation at a moment in time, and then tracing how and why the relation shifts over time. This focus on description and explanation can be subdivided into two central questions: (1) What causes the mean value of a trait in a population? and (2) What causes individual variation between people? These two questions are matched to corresponding methodologies.

Once a question is agreed on, the agreement often leads to a core set of methodologies (Overton, 2006). For example, Eisenberg (Eisenberg, 2000; Eisenberg, Champion, & Ma, 2004; Eisenberg & Spinrad, 2004) has carried out an extensive and highly influential line of research examining socioemotional development. Based on this line of work, she and her colleagues (Morris, Robinson, & Eisenberg, 2006) discussed the importance of using a multi-method approach to this work. These methods include self-report, informant report, direct behavioral observation, and biological correlates of emotion. These methods then can be used as part of a longitudinal or cross-sectional study. At the outset, however, Morris et al. (2006) noted that they would make little mention of either experimental methods or individual differences in their review because they studied socioemotional development. Readers interested in experimental methods were directed to the chapter in the volume reviewing social psychology (Smith & Harris, 2006) or reviews of cognitive development methods (Damon, Kuhn, & Siegler, 1998).

The methodological divide can reflect the researchers' core inquiry centered on Can versus Does. The first question asks if a specific mechanism or factor can shape the outcome of interest. For example, if we modulate levels of anger, can we see variation in levels of aggression? To ask this question, we often rely on experimental methods that manipulate a potential mechanism of interest and then carefully track any and all changes in the outcome. This is a mechanistic approach to the developmental question. However, this approach highlights the possible—not the probable. The other side of the equation looks to see if this mechanism actually does play a role in development under normal conditions. For example, do variations in levels of anger among children link to more aggressive behavior? For this work, researchers often call on observations in relatively naturalistic settings. An example of this dual question can be seen in classic work in the literature examining the development of anger and aggression in children (Coie & Dodge, 1998).

Observational studies have examined the contexts in which children show aggression (Maccoby, 1998), the general characteristics of aggressive children (Fry & Gabriel, 1994), and the environmental correlates of aggression (Andison, 1977). Acts of aggression then are tied to markers for anger. This work shows that anger, and subsequent aggression, can be triggered by limited resources; that overt aggression (often erroneously seen as a proxy for anger) is more often evident in boys; and that anger and aggression may be associated with exposure to exemplars, such as through viewing violent television programs. However, these conclusions are drawn from patterns of aggression (the operationalization of anger) that are clearly visible and tied to tangible triggers, underplaying more subtle acts of aggression that may arise from equivalent experiences of anger (Crick et al., 2001).

Systematically observing anger among children, even in the form of aggressive behavior, is difficult (Underwood, Galen, & Paquette, 2001) and often requires covert operations (Pepler & Craig, 1995). Experimental studies have long been used to both generate and explain instances of anger and aggression (Lewin, Lippitt, & White, 1939). In general, they build on available theory to isolate potential functional mechanisms, which may or may not be linked to presumed underlying levels of anger. For example, Cohen and Prinstein (2006) began with the theoretical assumption that adolescents wish to enhance their status among peers. This social goal, coupled with a belief that emulating a high-status peer will enhance their own standing, will lead adolescents to endorse risky or aggressive behavior they typically would reject if left to their own devices. For this test of the theory, Cohen and Prinstein selected adolescents deemed to have average levels of social status, based on peer nominations. This was a "control" factor to try to minimize the impact of participant characteristics (but see next paragraph). Participants then were led to believe that they were to interact via a chatroom with either high- or low-social-status peers from their school, again based on the same initial selection ratings. As expected, participants endorsed and engaged in aggressive social interactions when paired with an aggressive high-status peer relative to the low-status peer.

As a brief aside, we note that the aggression literature illustrates two of the core arguments made in this chapter. First, Cohen and Prinstein (2006) noted significant individual differences in the impact of their experimental manipulation. In particular, highly anxious adolescents were more susceptible to peer contagion relative to their less anxious peers. Throughout the chapter, we return to the argument that strong individual differences often are found in the socioemotional literature, and they are an important conduit to better understanding our constructs of interest. Second, this literature also illustrates the links between naturalistic observation and experimental manipulation within the field. As a case in point, Anderson and Bushman (1997) conducted a meta-analysis to examine the relation between aggression in the laboratory and aggression in context. They argued for a strong connection between the two sources of data. Of interest here is the title of their paper, "External Validity of 'Trivial' Experiments: The Case of Laboratory Aggression."

The literature on emotion and aggression captures how both mechanistic ("can") and observational ("do") studies provide unique information regarding socioemotional development. Longitudinal studies can help capture the rich descriptive architectures of development. Although they may lack the controls needed for strict causal inferences, new and evolving statistical techniques can capture this complexity (Khoo, West, Wu, & Kwok, 2006). These methods point to strong and plausible causal mechanisms that emerge in variable interrelations that shift over time.

However, given the intensive nature of these studies, trade-offs often are made. Researchers either focus on relatively smaller samples but institute rich and multilayered observation of development (N. A. Fox, Snidman, Haas, Degnan, & Kagan, 2015; Klein, Dyson, Kujawa, & Kotov, 2012), or they take a population approach, with less intensive measures that allow for more heterogeneous and variable samples (Chantala & Tabor, 2010). Longitudinal studies, by definition, are designed to capture a phenomenon over time. As such, they set at the outset methodologies and measures that often must be carried out without change within and across time. Shared measures allow for direct comparisons across the entire cohort at any one wave of study (T1) and over the life of the study (T1 versus T2 versus T3, etc.). Thus, they are a snapshot in the history of the field, reflecting the agreed-on outlines of knowledge at a point in time. Yet, as all developmental researchers have encountered, the data and theory supporting their carefully considered design can shift underneath their feet even as they carry out their studies. Science and time marches on, and scientists are then left to react.

Often short-term experimental studies can help probe the contours of new knowledge raised within longitudinal studies. The use of experimental studies may add to our shared understanding in a timelier manner than the ongoing longitudinal study. If we are lucky, we can add additional measures to later study waves in order to reflect the changing scientific consensus. Even when doing so is not possible, researchers can shift the questions asked and reconfigure study variables to reflect the new questions of interest. In this way, manipulations that create short-term variations in emotional state can be used to outline long-term trait-level profiles. Thus, new knowledge can emerge from a (partially) redundant network of studies, as described and illustrated (see Figure 9.1) six decades ago (Cronbach, 1957).

Current Approaches to Emotion Development

As noted, we approach emotions as biologically prepared adaptive processes that act in context for personal goal-setting (P. M. Cole

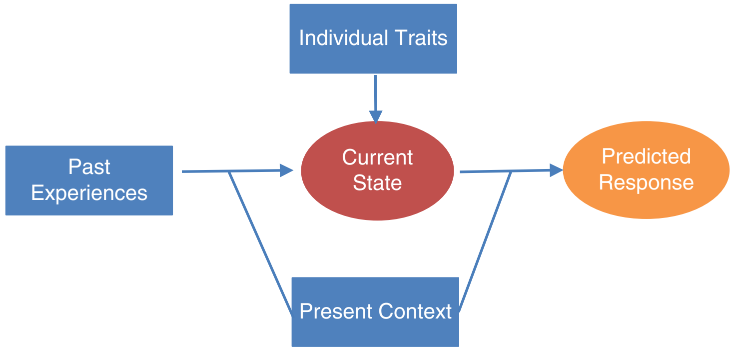


Figure 9.1 Illustrative model, based on Cronbach (1957), for a framework of processes thought to shape development. Overlapping sets of studies varying in design and timing that focus on subcomponents of the model can help improve our understanding of more complex and integrated systems. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

et al., 2004; Hastings, Miller, et al., 2014). From a developmental perspective, emotions are the initial language of communication (Tronick, 1989) and the core scaffolding for social interaction (Grossmann, 2015). Emotions are often in their rawest and purest form in infancy, before the emergence of stable and effective self-regulatory mechanisms. Given their pervasive role in development and functioning, it is not surprising that emotions, as a class, are complex and multifaceted. Emotions can vary by valence and intensity, the identity and context of probable triggers, their latency to emerge, the time to subsist, and the ease of regulation (Hastings, Kahle, & Han, 2014). This structural complexity is paralleled by equal complexity in the measures used to capture emotion. These measures each vary in chronometry, sensitivity to fluctuations in state, pace of change over the course of development, connection to social relationships, and level of association with biological functions, variations that have important adaptive survival processes that often are linked only tangentially to the socioemotional processes we as researchers wish to capture. In our discussion, we build from Larsen and Prizmic-Larsen (2006), who suggested that our measures of emotions generally have focused on three broad categories: language, behavior, and physiology.

Language measures provide an opportunity to capture an individual's subjective experience. For a field focused on affect, one could argue that this is '*the*' marker of our construct of interest. It seems intuitive that to measure anger, it is incumbent that we focus on individuals who are experiencing anger. We know a person is angry when they report "I feel angry." However, language measures are also often fraught with considerations above and beyond developmental change in the underlying emotion. Linguistic measures create the puzzle of disentangling changes in the construct to be reported (anger), the child's understanding of the construct (what are the typical antecedents and phenomenology of anger), and the ability (e.g., of preverbal infants) and willingness ("Will I get in trouble if I say I'm angry?") to report on their emotional states. Repeated probes across contexts or with clever interviewers (e.g., puppets; Measelle, Ablow, Cowan, & Cowan, 1998) can help reveal subjective states. For example, the Berkeley Puppet Interview (Measelle et al., 1998) was designed to assess children's self-concept but has been used to assess emotional and behavioral patterns in children as young as age 4. In this way, the method works to "meet" the child where he or she is developmentally—a general recommendation for both naturalistic and experimental methods.

Within the laboratory, researchers have attempted to manipulate task conditions to see if the pattern of report changes across context and time. For example, in the disappointing toy paradigm (Saarni, 1984), children are presented with a less-than-engaging present (e.g., a truck without a wheel; an unwound Slinky). By manipulating the presence and absence of additional people, the specific task demands, and the expectations built prior to the toy's presentation, researchers can note variation in report. Generally, children are less likely to report disappointment in the presence of an experimenter or parent (Tobin & Graziano, 2011) or when the experimenter attributes personal significance to the toy (Talwar, Murphy, & Lee, 2007). In this way, researchers can note and chart the emergence of children's understanding of—or ability to conform to—display rules, which govern when and how socially sanctioned emotions can be expressed (see Chapter 15 in this volume; Zeman & Garber, 1996). As noted in the discussion concerning anger, there are marked individual differences in this literature as well. In particular, there are sex (P. M. Cole, 1986), temperament (Kieras, Tobin, Graziano, & Rothbart, 2005), and culturally linked (Garrett-Peters & Fox, 2007; Yap, Ji, & Hong, 2017) variations in children's willingness to show displeasure within the paradigm.

Another way to work around a child's linguistic shortcomings is to rely on parental or teacher report (De Los Reyes, 2011; Morris et al., 2006). Parents and/or teachers have had the opportunity to interact with and observe the child across time and context. As such, they may have a more comprehensive view of core characteristics than can be captured either by the occasional observer or by laboratory manipulation. However, informant report also can add a level of complexity, as we are relying on the informant's ability to accurately assess and report on the target child's internal mental states. Of course, this added layer of subjectivity then means that it is now left to the researcher to infer the child's mental states underlying the adult's report—always a tricky proposition.

Although adult informants are chosen for their linguistic and cognitive sophistication, as well as their broader experience with the child, ironically, parents and teachers may be more open than children to the pressures of demand characteristics and measurement reactivity. Demand characteristics suggest that individuals are reluctant to endorse patterns of thought and behavior frowned on in the social context (Polivy & Doyle, 1980). Thus, one may find that questionnaire data produce children who are more prosocial and less aggressive than the children observed roaming the wilds of the playground (Underwood et al., 2001). This concern goes hand in hand with issues of measurement reactivity, which reflects the extent to which the methodology raises the participant's awareness of the construct of interest and consequently influences the participant's ability to modify behavior or response (Smith & Harris, 2006). Adults are much more likely to be alert to the "hidden" agenda behind our questioning and to shift their responses accordingly.

The second category of emotion research from Larsen and Prizmic-Larsen (2006) embraces the role of behavior. In many ways, behavior is the core concern of this chapter, as we have discussed the two competing ideals within socioemotional research. The first ideal calls for observing behavior in naturalistic settings, documenting developmental arcs as they emerge in response to new contexts and challenges. The second ideal targets specific contexts and processes to then manipulate and shift in order to observe and document subsequent behavior. We argue that these approaches fundamentally rely on each other to validate, expand, and apply the knowledge generated about emotional behavior.

Coupling observation and experimental manipulation may help us address core developmental trajectories in the experience and expression of emotion. For example, we may be interested in the prevalence and intensity of negative and positive affect in children and the subjective impact of peer interactions. Early in life, it may be easier to capture these socioemotional constructs in a natural setting, as young children often are expressive, boisterous, and loud. We are less likely to see the same behaviors and interactions as clearly in older children and adolescents, particularly as children develop more self-regulatory skills (Pérez-Edgar, 2015) that will be carried into adulthood (see Chapter 15 in this volume), reflecting display rules shaped by the larger society and their specific peer networks (Zeman & Garber, 1996).

Thus, research in a lab setting, even without an experimental manipulation, may increase our ability to see our constructs of interest as they become more internalized and are communicated more subtly. Indeed, researchers often move a naturalistic observation into the laboratory in order to capture more closely socioemotional processes during set events. For example, Eisenberg and colleagues have had parents and adolescents discuss issues of concern in order to closely observe patterns of emotion during the conversation (e.g., Hofer et al., 2013). Luckily, the laboratory setting then opens up the opportunity for experimental manipulation as well. Potential manipulations can include shifting the discussion partners across family members (McDowell, Kim, O'Neil, & Parke, 2002), priming specific emotions (Jouriles, Murphy, & O'Leary, 1989), or modifying the difficulty of a shared task (Dennis, 2006). The laboratory setting also opens up the opportunity for new, previously unavailable, levels of analysis.

The third broad category of research (Larsen & Prizmic-Larsen, 2006) encompasses the use of physiological and neural measures to capture emotion processes embedded in biological systems. This is the newest addition to the research arsenal for affective developmental science. Its relative novelty, complexity, and cost have led to a vague sense that bio-based measures are a special class of data onto itself. That is, there is a tendency to treat evidence of neural patterns associated with emotion as prima facie evidence for a "mechanism." However, mechanisms arise from functional influences on current state and lawfully direct change over time. As van der Molen and Molenaar (1994) noted, "[T]he usefulness of psychophysiological measures depends on the demonstration of the sensitivity of the measures to task manipulations derived from developmental psychology" (p. 466). Neural activity can act as a descriptive or correlational data point to our end state of interest (Gee et al., 2013; Giedd et al., 1999) as easily as it can serve as the conduit for change within an experimental study (McClure et al., 2007; Pérez-Edgar et al., 2007). Finally, these measures can spur on and reflect emotions, acting as both the mechanism and the embodiment of underlying change. Thus, method and theory are just as important when it comes to interpreting bio-based measures as when we examine linguistic or behavioral data.

Physiological and neural measures also have distinct advantages in our attempt to capture emotion-linked processes over time. First, as long as the measure is tolerable to children, we can characterize the same measure, in the same form, repeatedly over time (Wolfe & Bell, 2007). With linguistic and behavioral data, developmental psychologists must carefully titrate and change measures in order to meet the child's current level of functioning and typical contexts. Doing so often places the developmental psychologist at odds with the statistician, who may insist that change over time can be captured only when the identical measure is at hand at each time point (Willett, Singer, & Martin, 1998). Although we all benefit from recent statistical advances sensitive to developmental considerations (Foster, 2010), biological measures (at least in their surface presentation) could show the needed mark of homotypic continuity, minimizing these analytic concerns. As such, they can contribute to our attempt to capture emotional processes by revealing patterns of stability (or change) that otherwise would be obscured by development in higher-order markers of emotion. The temperament literature has made extensive use of electroencephalogram activity from infancy through adolescence to reveal patterns of approach and avoidance motivations tied to general affective patterns as well as individual differences in trait-level markers and individual response to specific affective challenges (N. A. Fox, Hane, & Pérez-Edgar, 2006; N. A. Fox, Kirwan, & Reeb-Sutherland, 2012). In another example, Forbes, Fox, Cohn, Galles, and Kovacs (2006) found that children in the disappointing toy paradigm show a physiological response to receiving the toy, even in the absence of overt facial changes and report. Thus, the use of this task over time can help researchers disassociate the developmental arc of emotion expression from underlying markers of emotional reactivity.

In addition, the temporal dynamics of emotion can be better captured through the use of psychophysiology, often integrated with behavioral and linguistic measures (N. A. Fox et al., 2012). When bringing together our populations of interest (often children) with our specific content questions, this new subniche of research falls under the umbrella of developmental affective psychophysiology (Hastings, Kahle, & Han, 2014). Theoretically, one would expect that a triggering event or thought will create a temporal cascade across levels of emotion. Based on information processing models (Dodge, 1991), the child begins by encoding and interpreting emotion-linked cues. From there, the succession of goal clarification, response construction, and response selection occur. Accompanying these fundamentally cognitive processes, there are also behavioral mechanisms that ebb and flow at each point in time. These cascades are internal to the child and are evident only later in overt behavior or report. These "outcomes" cannot track the pattern of emotion, as they reduce all contributing processes to a single overt marker. This blurring of multiple processes can obscure any true signals of emotion. Thus, we are challenged to align the chronometry of emotion with the chronometry of our measures. Some methods, such as questionnaires, cannot adequately measure chronometry; they must work at the level of chronicity. For that reason, multiple strategies are needed to better capture the emotion patterns seen within and across individuals. In particular, we need measures that are temporally and spatially sensitive, allowing researchers to better delineate the components and time course of emotion processes.

Biological measures reflect in real time a complex and multilayered web of mechanisms that draw on processes that emerge from the brain and encompass the entire body. For example, central nervous system measures have examined brain structure (diffusion tensor imaging) and function (functional magnetic resonance imaging), both at rest and in response to affective stimuli. Emotion-linked neural activity is distributed widely across the brain, reflecting the pervasive and flexible role of emotion in cognition and behavior (Hastings, Kahle, et al., 2014). For example, the amygdala, striatum, and prefrontal cortex are implicated as hubs for core functions of emotion triggers, including punishment, reward, and self-monitoring. The systematic coordination of activity is central to adaptive and flexible functioning.

A series of studies in the temperament literature has shown that patterns of responses in limbic, striatal, and prefrontal regions are associated with patterns of negative and positive affect, from childhood into adulthood (N. A. Fox et al., 2012; N. A. Fox & Helfinstein, 2013). Much of the translational interest in behavioral inhibition is based on its association with increased risk for social anxiety disorder (Clauss & Blackford, 2012). Although the clinical literature historically has been a separate line of research from developmental science, this work also has found a distributed neural network associated with pediatric anxiety, encompassing the amygdala, the bed nucleus of the stria terminalis, the striatum, and multiple regions of the prefrontal cortex (Blackford & Pine, 2012).

The first major study of amygdala function in behavioral inhibition found that young adults categorized as inhibited in the second year of life showed significant bilateral amygdalar activation to the presentation of novel faces versus fixation relative to participants without a history of behavioral inhibition (Schwartz, Wright, Shin, Kagan, & Rauch, 2003). Fox and colleagues (Pérez-Edgar et al., 2007) found that adolescents with a history of childhood behavioral inhibition also showed increased amygdala activation when attending to their subjective fear of emotion faces during an attention-emotion face task, coupled with amygdala deactivation when passively viewing the same faces. Intriguingly, McClure and colleagues (2007) found the same pattern of hyperactivation and deactivation across conditions in clinically anxious adolescents completing the identical functional magnetic resonance imaging task.

Guyer et al. (2006) were among the first to show that adolescents with a history of behavioral inhibition showed increased striatal response in anticipation of increasing monetary rewards. Building on this initial finding, it may be that the striatal hyperresponse to monetary reward is evident only when reward is contingent on one's own performance (Bar-Haim et al., 2009) and when anticipated feedback is met with a negative response (Helfinstein et al., 2011). This heightened striatal response in behavioral inhibition also is linked to increased anxiety, particularly in the context of genetic risk (Pérez-Edgar, Hardee, et al., 2014) and increased levels of substance use (Lahat et al., 2012). Adolescents with a history of behavioral inhibition also show striatal hypersensitivity to anticipated social evaluation (Guyer et al., 2014), particularly if previously exposed to harsh parenting (Guyer et al., 2015).

To fully influence emotional functioning, the central nervous system must work to shape how the rest of the bodily systems help the child move through and react to developmental challenges. The autonomic nervous system regulates the functions of our internal organs and acts as a bidirectional conduit for responding, assessing, and exploiting the environment (Hastings, Kahle, et al., 2014). Within the umbrella of the autonomic nervous system, the sympathetic (SNS) and parasympathetic systems work together to carry out allodynamic control (Hastings, Kahle, et al., 2014), carefully titrating the individual's internal experiences and how the person navigates the environment.

For example, young children were asked to carry out an impossible task designed to trigger frustration and other negative emotions (Kahle, Miller, Lopez, & Hastings, 2016). They were asked to draw a "perfect" green circle and then repeatedly told they failed to meet this ideal. Across the 3-minute task, a measure of SNS (pre-ejection period, PEP) was assessed 10 times. Researchers then measured PEP three times during a 1-minute recovery period in which the child was praised for the last circle drawn. Across this time series, Kahle et al. found that there was shortening PEP (SNS activation) during the frustration induction, followed by lengthening PEP (SNS inhibition) during the recovery period. Further, children who had more SNS activation during the induction were observed to express more anger during the recovery period. Thus, this SNS pattern may reflect the chronometry of one physiological aspect of the onset, experience, and attenuation of an anger response to the laboratory-induced frustration, with a more persistent behavioral aspect of the anger experience being evident downstream in the time course of the emotion.

Within the behavioral inhibition literature, autonomic nervous system measures have proved crucial to examining the proposed mechanisms underlying the phenotype. The complex behavioral pattern seen in behavioral inhibition is hypothesized to result from a hyperaroused limbic system, centered on the amygdala. In proposing this model, Kagan, Reznick, and Snidman (1988) drew on a line of research linking the amygdala to the acquisition of conditioned fear (Davis, Walker, & Lee, 1997), the induction of vigorous limb movements (Amaral, Price, Pitkanen, & Carmichael, 1992), and the modulation of distress cries (Newman, 1985). However, there were (and are) methodological and developmental barriers to examining the functioning of amygdala directly in children across age and context. Thus, two decades of studies chose behavioral or physiological outcomes that reflected the presumed activation of nuclei within the amygdala.

For example, children and adolescents with a history of behavioral inhibition showed elevated attention to novelty (Marshall, Reeb, & Fox, 2009; Reeb-Sutherland, Vanderwert, et al., 2009) and to behavioral errors (McDermott et al., 2009) and had difficulty disengaging from threat cues (Pérez-Edgar & Fox, 2005), as marked by event-related potentials. They also exhibited increased potentiated startle to threat (Reeb-Sutherland, Helfinstein, et al., 2009), greater right frontal electroencephalogram asymmetry at rest (Calkins, Fox, & Marshall, 1996; Hane, Fox, Henderson, & Marshall, 2008), perturbations in salivary cortisol levels at rest and after provocation (Pérez-Edgar, Schmidt, Henderson, Schulkin, & Fox, 2008; Schmidt, Fox, Sternberg, et al., 1999), higher heart rates and lower heart rate variability at rest (Marshall & Stevenson-Hinde, 1998), unique patterns of cardiac reactivity to emotionally laden narratives (Bar-Haim, Fox, VanMeenen, & Marshall, 2004), and lower thresholds for detecting threat in the environment (LoBue & Pérez-Edgar, 2014; Reeb-Sutherland et al., 2015). In each case, the specific functional marker was chosen for study because it was hypothesized to grow out of the hypersensitive limbic response to novelty and uncertainty presumed to fuel the observed behavioral profile of behavioral inhibition (Kagan, 2012; White, Lamm, Helfinstein, & Fox, 2012).

Although the biological system is carefully interwoven in functioning, our methods are not always adequate to unlocking these mechanisms (Miskovic & Schmidt, 2012; Schmidt & Segalowitz, 2008). Indeed, Larsen and Prizmic-Larsen (2006) suggested that "the various components of emotion will never correlate substantially with each other" (p. 342). For example, Nesse et al. (1985) examined measures of distress during in vivo exposure therapy in phobic individuals. Although they noted increases in subjective anxiety, pulse, blood pressure, plasma norepinephrine, epinephrine, insulin, cortisol, and growth hormone, there was only modest convergence in the "magnitude, consistency, timing, and concordance" (p. 320) of their measures.

Our empirical imprecision reflects the fact that the response systems we use as core measures of emotion have roles that extend far beyond serving as indicators of emotion. Although cardiac activity provides the foundation for calculations of heart rate, heart rate variability, and respiratory sinus arrhythmia, these relatively discrete markers of emotion and mood state are embedded within much broader and pervasive survival functions (Dennis et al., 2012; Hastings, Kahle, et al., 2014). These markers will function separately from, and independent of, our interests in capturing variations in affective development and emotional functioning. Of course, there is the added difficulty that the biomarkers of emotion also mutually influence each other, and each has its own developmental trajectory. Although infants may show a robust startle response from birth (N. A. Fox et al., 2012; Marshall et al., 2009), studies examining electrophysiological measures have to take into account a slow developmental progression in form and function (Bell & Cuevas, 2012; Bell & Wolfe, 2007; Cuevas & Bell, 2011). Thus, we need close observation in the laboratory, ideally over time, to isolate and track fluctuations that follow (even if weakly) our experimental manipulations. Here, we also rely on newly emerging statistical tools that can handle multiple weakly correlated measures.

One final item to note regarding the categories of functioning laid out by Larsen and Prizmic-Larsen (2006) is that the authors do not explicitly make mention of cognition, and cognitive processes, in measuring emotion. This, in part, reflects two historical forces. First, as mentioned earlier, is the notion that socioemotional development should rely on naturalistic observation while cognitive processes are best understood through careful manipulation in the laboratory. This methodological distinction also reflects a second historical force—namely, the assumption that cognitive and emotional processes are qualitatively different, playing out across different realms of complexity and control and reliant on different underlying neural substrates. From this perspective, emotions are visceral and somewhat primitive states that emerge early in development and are controlled only loosely. Cognitive processes, in contrast, emerge later in development, are controlled and deliberate, and work externally to impose order on emotions run amok.

However, this view underestimates both the emergence of cognitive competencies in infants and the degree of integration between cognition and emotion across development. Recent work on affect-biased attention suggests that core cognitive mechanisms may play a core role in socioemotional development (Pérez-Edgar, Taber-Thomas, Auday, & Morales, 2014). Affect-biased attention, as used by Todd, Cunningham, Anderson, and Thompson (2012), refers to "attentional biases that cause preferential perception of [any] particular category of stimulus based on its relative affective salience" (p. 365). It is likely that affect-biased attention influences cognitive and emotional development from infancy (Morales, Fu, & Pérez-Edgar, 2016). For example, preferential attention allocation toward emotionally salient objects emerges early in development, likely due to specific perceptual markers (LoBue, Rakison, & DeLoache, 2010). In the competition for limited attentional resources, infants prioritize objects that decrease danger and increase reward (Peltola, Leppänen, Palokangas, & Hietanen, 2008).

No other object is as closely tied to survival, punishment, and reward as the human face (Hoehl & Striano, 2010). Due to the coupling of perceptual cues, rewarding daily events (e.g., feeding), and long hours of exposure, infants quickly begin to show preferential looking to human faces (Leppänen & Nelson, 2009). This preference is magnified when the face also conveys an emotional threat signal. Thus, affect-biased attention is early appearing, likely rooted in evolutionary concerns, and has the potential to influence broad patterns of socioemotional behavior throughout life. Affect-biased attention, particularly if stable and entrenched, also can act as a developmental tether that helps sustain early socioemotional and behavioral profiles over time, even in the face of internal and external forces that typically act to shape early tendencies (Pérez-Edgar, Taber-Thomas, et al., 2014). Testing this argument opens another door to the use of experimental methods within developmental affective science.

BRIDGING TWO RESEARCH TRADITIONS

Experimental Work in Emotion Development

Although researchers enter into heated debates over the nature of emotion, characterizing them as discrete entities (Izard, 1993) or varying across spectrums (Barrett, 1998), there is general agreement that emotions are multifaceted, expressed at multiple levels of functioning, and subserved by processes that emerge and recede over the course of time and across contexts. The thorny question then becomes how to best (try to) tame each of these components, "slowing" them down just enough so that we can capture them briefly and track their paths. In this sense, the emotion researcher is the kindred spirit of the particle physicist.

Morris and colleagues (2006) argued that developmental psychology is somewhat unique relative to other subspecialties in its strong emphasis on context. Behavior, thought, and change in behavior and thought are believed to be highly influenced by, and perhaps limited to, specific contexts (Anderson & Bushman, 1997). Given this point of view, experimental methods, particularly if in the laboratory, are thought to be divorced from context. The laboratory setting is thought to sacrifice the probable ("does") in favor of the potential ("'can'"). The overarching argument of the chapter has been that experimental methods are not at odds with descriptive/correlational work. Rather, these two streams of science are complementary and, indeed, depend on each other for advancement.

Specific strengths to the experimental approach complement the strengths of nonexperimental methods. In experiments, (1) we manipulate the independent variable, (2) participants are sorted via random assignment, and (3) we have control over the operation of the variables and the general setting. These three traits, as a set, allow researchers to focus specifically on questions of interest while controlling, or pushing to the background, factors known (and unknown) that also may be at play in our specific dependent variable. This experimental meta-method provides the traditional foundation for being able to infer the presence of causal relations between variables.

A focus on manipulation naturally constrains the level of analysis (Loeber & Farrington, 1994). Researchers can quickly, and rightly, point to pivotal constructs at the heart of emotion development that cannot be manipulated in a "true" experimental fashion. These constructs include theoretical titans, such as temperament, socioeconomic status, ethnicity, and gender. Foundational work has shown how each of these factors covary with behavior, cognition, and adaptive outcome across development (M. Lewis, 2010). In addition, emotions are generated by complex social and contextual processes that are not readily replicable in a laboratory setting. However, this is not to say that nonrandomizable traits cannot be examined productively through experimental methods. Rather, we are emphasizing that we can use identified traits to see how individuals respond to variations in controlled stimuli and situations.

For example, hormonal systems are central to the individual's ability to respond to environmental challenges quickly and flexibly. Marceau and colleagues (2014) examined within-person coupling across cortisol, dehydroepiandrosterone, and testosterone across three conditions: parent–adolescent conflict discussion (anger induction), social performance (anxiety induction), and venipuncture (pain and fear induction). They found that, unlike in adults, adolescents show positive coupling across all three stressors. Layering on individual difference factors, Han, Miller, Cole, Zahn-Waxler, and Hastings (2015) found that adolescents with more externalizing problems had greater positive coupling between cortisol and testosterone in the context of the conflict discussion. Manipulating context to (somewhat specifically) elicit anger, therefore, allowed for the identification of relations between trait-level characteristics and hormonal systems that could not be manipulated directly themselves. Yet lacking random assignment, this could not be considered a true experiment.

Another factor that may disadvantage the experimental approach is our core belief that development is reflected in change over time, often associated with age (Overton, 2006). Traditionally, the argument has been that one cannot capture developmental change, *true* change, in the experimental laboratory. Capturing change opens the window to allow researchers to look for, or unearth, the mechanisms that support developmental trajectories. However, age, in and of itself, is not a developmental mechanism. It is a convenient time marker. Thus, we need clever and systematic studies that observe processes that covary with constructs of interest, coupled with direct manipulation whenever possible. Carefully designed and focused experimental studies can verify the nature of the relations that are identified in descriptive cross-sectional and longitudinal work.

For example, with age there are marked changes in the expressive function of an affective signal. Babies can cry for multiple reasons. They are hungry, tired, ill, angry, or simply bored. Cries, over time, take on greater specificity and are tailored more narrowly to specific triggers. With time, the crying may cease altogether, replaced by distinct vocalizations or regulated into a fully internalized form. It is unlikely that the variation in the presence of a putative cry trigger and subsequent response needed for systematic study would emerge spontaneously in a natural setting. Experimental studies have worked to capture this variation by building on reliable laboratory-based cognitive tasks and layering on an affective component by varying either the stimulus or the testing context. One example of such an approach is the Laboratory Assessment Battery of Temperament, which has a set sequence of interactions and objects designed to elicit a range of emotions in toddlers and preschool-age children, including anger, frustration, sadness, and joy (Buss & Goldsmith, 2000).

For older children, we also can manipulate standardized cognitive tasks in order to modulate the emotional content, or context, of task performance (Prencipe et al., 2011; Zelazo, Qu, & Müller, 2005). Traditional task variants are designated as "cool" while affectively charged variants are "hot." Performance in the cool and hot conditions is then compared to isolate the impact of an emotional component on the cognitive, behavioral, or physiological measure of interest.

For example, Lewis and colleagues (M. D. Lewis, Hitchcock, & Sullivan, 2004; M. D. Lewis, Lamm, Segalowitz, Stieben, & Zelazo, 2006; M. D. Lewis & Stieben, 2004) have carried out a series of studies using variants of the standard go/no-go task. In this task, children are instructed to make a response when they view a stimulus (e.g., an X) and refrain from responding when presented with another stimulus (e.g., an O). By modulating the ratio of go and no-go trials, researchers can shift accuracy rates and reaction times. Performance on the standard task is thought to reflect development in frontostriatal networks (Durston et al., 2002). In the affective version of the task, Lewis found that negative mood induction increased electrophysiological responses to the task, relative to the cool baseline condition (M. D. Lewis & Stieben, 2004).

The study of emotional development is complicated further by the fact that our construct of interest, emotion, is coupled very quickly to a process, emotion regulation, designed to push and pull the initial construct. (See Chapter 15 in this volume.) Thus, the psychologist must be nimble in the attempt to measure a construct even as the object of study is employing *that very construct* to modify its behavioral manifestation. This is a thorny issue for all developmental psychologists, but we suggest that experimental methods may help us puzzle it through. Standardized laboratory tasks can help track if the responses triggered by the same stimulus/condition/manipulation change across development.

For example, a number of empirical tasks have been designed to assess the interplay between emotion and effortful control. These include the Stroop color-word task (Stroop, 1935) and its emotional variants (Pérez-Edgar & Fox, 2003), the Stroop-analog Day-Night task (Gerstadt, Hong, & Diamond, 1994) and its emotional Happy-Sad variant (Lagattuta, Sayfan, & Monsour, 2011), the go/no-go task (Casey et al., 1997), the spatial conflict task (Gerardi-Caulton, 2000), and the flanker task (Eriksen, 1995). Other than the original Stroop task, none of these tasks requires reading competency. All of the tasks also can be designed for use with adults and are amenable for use with psychophysiological and imaging techniques (N. A. Fox et al., 2006). Each task has been modified to meet the skill and interest level of children, from toddlerhood through adolescence (Pérez-Edgar & Bar-Haim, 2010). Since children generally become progressively better at masking their affective or cognitive responses to our laboratory tasks, in vivo methods are particularly useful in revealing underlying neurobiological patterns of reactivity and regulation (Luna & Sweeney, 2004).

Finally, we suggest that experimental methods may be particularly beneficial in charting emotion development from its nascent form to its adult manifestation. Cognitive development often explicitly looks to note when and how children come to acquire a specific skill or come to match the final adult form. This focus on timing and mechanism may be due to the fact that researchers regularly deal with constructs that can be placed on concrete and readily agreed-on metrics. Thus, there are a multitude of studies tracking the developmental pathways of measures, such as vocabulary size, numerosity, and spatial reasoning.

The same perspective can be harder to impose on socioemotional measures. When simply focusing on a construct in isolation, it can be difficult to state whether a child is "better" or "worse" in expressing anger or happiness. This relative argument does not make a great deal of sense. It may be better to examine when, how, or if a child can engage and manage emotion in order to reliably and effectively attain a specific goal. Thus, the comparative metric is not built into the emotion per se but into the functional impact of the emotion on broader patterns of behavior. In the same vein, emotion regulation, when effective, increases the degree to which an emotion facilitates attaining a goal. The mark of regulation or dysregulation may be specific to the goal at hand as well as the context in which the child is embedded (Hastings, Kahle, et al., 2014). Then we can step back and see if the underlying mechanism for this affective pattern, and the context in which is evoked, is the same. Emotions often are "felt" below the surface, but not overtly expressed, perhaps, ironically, due to the development of self-regulation. Neural and electrophysiological mechanisms can play an important role in getting "under the skin."

Thus, the use of multiple levels of analysis is crucial. However, experimental work has not always embraced this challenge. Morris et al. (2006) criticized experimental designs in that "these types of studies rarely use a multimethod approach, instead relying primarily on the experimental tasks to assess a specified cognitive ability" (p. 379). This is not an unfair critique. However, the fault lies not in the method but in us, the researchers. Luckily, a focus on socioemotional functioning, especially in an era with more widespread access to bio-based measures (Dennis et al., 2012; Hastings, Kahle, et al., 2014), has pushed the field to embrace multiple methods, even at the cost of "messier" data than one would find with the traditional, tightly controlled, one independent variable: one dependent variable ideal.

Also creating noise and 'messiness' is the fact that child-centered nonrandomizable traits can, and often do, impact the data. Indeed, often there will be moments when individual differences overwhelm central tendencies (Hastings, Kahle, et al., 2014). Thus, we argue that both natural observation and experimental studies can benefit by embracing the specific traits and proclivities that children carry within themselves through their environments and into the laboratory.

Individual Differences in Emotion Development

If anything can unite the methodological and theoretical traditions of descriptive and experimental research discussed in this current chapter, it is the fact that neither camp is particularly excited by the presence of individual variation. For the traditional experimentalist, individual differences are an annoyance—an "error variance" that is to be eliminated through strict control at each level of the laboratory protocol. Although there is a role for equifinality and multifinality in our theoretical language, for the developmental psychologist, they create another mechanism and factor that needs to be chased down when evident in the data. Noting factors that impact trajectories is no small feat when dealing with either long-term careful observation or carefully controlled experimental studies manipulating a curated set of variables. There are inherent tensions between outlining nomothetic laws that focus on universal sequences and their contexts and identifying idiographic patterns that are unique to individuals (Scarr, 1992; Scarr & McCartney, 1983). Although both methodological camps are focused on the environment, either as a variable to control or as a factor to describe carefully, there is often the implicit assumption that the shape and meaning of any one environment can be presumed to be static within and across participants.

However, the environment does not have the same meaning for all individuals. Scarr (1992) argued that a child constructs a unique reality for him- or herself. Thus, individual differences are to be expected in any study of complex developmental functions. Within the temperament literature, one can see clear differences in how children react to ostensibly identical social contexts. Some children rush to embrace the novelty of the social world, while others pull back from ambiguous and unexpected threats. These variations appear early and shape the child's "experienced environment." In this way, fairly subtle individual differences can impact socioemotional functioning from infancy by creating cascading and self-reinforcing biases in social cognition and behavior. The current discussion has returned repeatedly to the importance of context for both emotion and emotional development. The focus on context is in line with the arguments that (1) emotions have broad developmental functions and (2) the success of any affective strategy is tied to the specific constraints and expectations of the environment. Recognizing that a given context—even an experimentally manipulated one—is not experienced in the same way by all people adds yet another level of complexity but also affords novel opportunities for insight. Petrill and Brody (2002) argued that experimental psychology creates variability by manipulating the environment, while researchers interested in individual differences study variations that occur naturally. In particular, the latter is focused on using statistical methods to "partition sources of variance in a measure." Bringing together both approaches may be particularly helpful when examining issues of socioemotional development.

For both cognitive and socioemotional studies, the most common and obvious individual difference factor seen in the literature is sex or gender. There is a long literature suggesting that boys and girls differ in the rate, intensity, and context under which they display positive and negative emotions (Brody, 1985; Kring & Gordon, 1998). However, much of this literature relies on the questionnaire report, from either child or an outside observer (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). Although these data are important for capturing subjective experience and broad patterns of response, they are hard-pressed to capture affect in the moment. In addition, observer report may be influenced by cultural expectations shaping acceptable presentation patterns in boys and girls (Chaplin, Cole, & Zahn-Waxler, 2005).

To examine the influence of methodology, Chaplin and Aldao (2013) examined patterns of emotion generated by direct observations of children. They found that gender-linked differences in affect and internalizing/externalizing difficulties emerged with age. There were no observed differences in infancy, with the first deviations evident in toddlerhood before widening in childhood.

In addition, the social context played a role, with sex-linked variation in emotional expression when alone versus when with parents and peers. This finding may reflect the internalization of display rules over time as a function of parental and cultural socialization (Klimes-Dougan et al., 2007; Zahn-Waxler et al., 2008). Alternately, it may be that biologically based differences in emotion expression may emerge over time due to the developmental trajectory of underlying mechanisms, such as the interplay between subcortical and prefrontal hyperreactivity (Hare et al., 2008). In distinguishing these possible mechanisms, we also can look to see if girls are better able to modulate their emotions or if they are more socially motivated to do so. Of course, it is important to point out that it would be difficult to gather the data needed to do so outside of a laboratory setting.

With respect to gender and other factors of interest, we are lucky, as a science, that individual differences are likely to be lawful rather than a random assortment of disconnected and independent traits. To capture these individual differences, we often have to shift from a focus on variance across conditions to variance among individuals. This change in focus is then coupled by a shift from a variable-centered analytic approach to a person-centered approach. Thus, the focus is not on how a variable behaves across context or time but on how individuals, or groups of individuals, navigate these environments. An example of this work can be found in the pediatric clinical literature.

Careful observational studies have noted that a subset of children show elevated levels of mood dysregulation and irritability that are extreme, long lasting, and outside developmentally appropriate norms (Leibenluft, 2011). Mood dysregulation also shows a phenomenological and diagnostic link to bipolar disorder (Leibenluft & Rich, 2008). However, there is an ongoing debate as to whether the observed behavior is a developmental precursor to the adult form of bipolar disorder (a prodrome) or a disorder onto itself (e.g., disruptive mood dysregulation disorder from the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders*). To help disentangle these associated constructs, Rich and colleagues (2005, 2007) completed a set of experimental studies using the Affective Posner task.

In the Posner cued-attention task (Posner & Cohen, 1984), participants are presented with a single cue on one side of the visual field (left or right). The cue is meant to pull attention to its location automatically. Following cue removal, a target probe appears either in the same location as the cue (valid trials) or on the opposite side of the cue (invalid trials). Participants are asked to press a button to indicate probe location. The validity score is the reaction time on invalid trials minus valid trials. This difference (the validity effect) is believed to represent the effort required to disengage from the cue location in order to shift to the probe location. In the affective Posner task, the cues have an emotional valence based on punishment and reward cues, emotional words, or emotional facial expressions (Derryberry & Reed, 2002; E. Fox, Russo, Bowles, & Dutton, 2001; Pérez-Edgar, Fox, Cohn, & Kovacs, 2006).

In Rich's studies (Rich et al., 2005, 2007), the trial blocks all involved the same stimuli and task demands but differed in the contingencies presented with variations in performance. Block 1 served as the baseline, with children informed of the accuracy of their responses ("Good job!" or "Incorrect!"). In block 2, children won or lost 10 cents on the basis of their performance and were informed of the accuracy of their response and whether they had won or lost money. During block 3, correct responses resulted in accurate feedback and reward on 44% of trials, but on 56% of trials, rigged feedback informing participants that they had been too slow was provided randomly regardless of performance, and participants steadily lost money.

Using the experimental manipulation, with an eye to individual differences, the researchers found no difference at baseline between typically developing children, children with the narrowly defined bipolar phenotype, and children with severe mood dysregulation. However, as task demands increased, the typically developing children showed different response patterns relative to both risk groups. Using behavioral, neurophysiological (event-related potentials, and functional magnetic resonance imaging measures (Deveney et al., 2013)), this line of research has distinguished differential patterns of arousal with the titration of negative feedback and frustration.

As we have noted, large-scale longitudinal studies have shown a strong link between early behavioral inhibition and later social anxiety (Clauss & Blackford, 2012). Smaller-scale laboratory studies have been used in a complementary process to probe the mechanisms that potentially may underscore developmental patterns observed over time. For example, Pérez-Edgar and Fox (2005) had behaviorally inhibited children complete an affective Posner task similar to the one used with children with mood dysregulation (Rich et al., 2007). They found that compared to the traditional (affect-neutral) Posner task, performance in the affective Posner task was marked by dramatic decreases in reaction times, an increase in errors, an increased validity effect, and increased electrocortical activity. Temperamentally shy children in the study differed from their nonshy peers within the affective Posner task only. In addition, shy children preferentially attended to the negative cues presented during the task. Here the use of a controlled task setting provided the opportunity to compare performance across shared circumstances and contexts, and levels of analyses, that cannot be readily observed in more uncontrolled, natural settings.

CONCLUSION AND FUTURE DIRECTIONS

The current discussion has focused on historical tensions that have worked to shape the interplay between broad areas of research within developmental psychology. As Overton (2006) noted, an adherence to specific theoretical worldviews can result in downstream clashes regarding the methods used to study questions of interest. In many ways, methodology has played an unacknowledged role in aiding, or blocking, our willingness to cross empirical boundaries.

Some of these barriers initially appear trivial. For example, Smith and Harris (2006) noted that experimental studies within social psychology are focused on controlling, shifting, and tracking the independent variable. The ability to systematically target the independent variable is a core measure of scientific rigor. In contrast, developmental psychologists, particularly within naturalistic studies, often are devoted to capturing the dependent variable and then characterizing the factors that surround it. These differences make it a bit harder for researchers to build on each other's work, which adds to bifurcations in literatures and even less likelihood of cross-fertilization.

We continue to see these divisions today. For example, within developmental affective science, there is now a subdivision for developmental affective neuroscience. Within that world, there are questions regarding empirical links and boundaries to developmental *social* neuroscience and developmental *cognitive* neuroscience. However, even with these tendencies to define and divide, there exist current examples of (slow-moving) lines of research that draw on multiple traditions of naturalistic observation and experimental manipulations and incorporate both universal patterns and individual differences.

For example, our discussion of affect-biased attention is part of a larger literature suggesting that attention to threat may play a causal role in the emergence of clinical anxiety problems (Morales et al., 2016; Todd et al., 2012). From the observational side, there is correlational evidence that individuals high in anxiety show attention biases to threat (Wilson & MacLeod, 2003). Much of this work is based on biases captured via computer tasks, most often the dot-probe task (Roy, Dennis, & Warner, 2015; Todd et al., 2012). In this task, participants see a pair of stimuli simultaneously, most often for 500 ms; one stimulus is emotionally salient (e.g., threatening), and the other is neutral (e.g., nonthreatening). A probe replaces one of the two stimuli. The individual is required to respond as accurately and as quickly as possible to the probe. An attentional bias toward emotional stimuli is inferred when participants preferentially attend to emotional cues, resulting in decreased reaction times to probes replacing the emotional stimuli compared to the neutral stimuli. For instance, in the original dot-probe task, MacLeod and Mathews (MacLeod, Mathews, & Tata, 1986; Mathews & MacLeod, 1985) found that anxious individuals were faster to respond to probes following threatening words. In contrast, control participants were faster to respond to probes following neutral words, exhibiting a bias away from threat (MacLeod et al., 1986).

This pattern is evident from young childhood through adulthood, reinforcing the argument for an etiological role in anxiety (Shechner et al., 2012). Children diagnosed with an anxiety disorder also were shown to display an attentional bias toward threat compared to nonanxious controls (Roy et al., 2008). In addition, temperamentally at-risk, but healthy, adolescents displayed the same bias pattern (Pérez-Edgar, Bar-Haim, McDermott, et al., 2010). Finally, the magnitude of attention bias has been found to predict levels of anxiety symptoms (Waters, Mogg, Bradley, & Pine, 2008), suggesting a relation across the anxiety dimension and in childhood as well as adulthood. Interestingly, even when group-level main effects of bias are not evident, bias pattern interacts with behavioral inhibition status to predict social anxiety and social withdrawal (C. Cole, Zapp, Fettig, & Pérez-Edgar, 2016; Morales, Pérez-Edgar, & Buss, 2015; Pérez-Edgar et al., 2010, 2011; White et al., 2017).

However, as we have already noted, strong patterns of covariation are necessary, but not sufficient, to infer causality. Thus, researchers have worked to complement these initial observations with mechanistic studies that manipulate attention to threat and examine subsequent changes in anxiety (Roy et al., 2015). The most persuasive evidence for this comes from experimental investigations, such as attention bias modification (ABM) studies (Bar-Haim, 2010; Eldar et al., 2008; Hakamata et al., 2010).

In ABM studies, experimental manipulations of the attentional bias (i.e., reducing or augmenting the bias) in children and adults are examined to see if they lead to the expected changes in anxious thought and behavior (i.e., reduction or augmentation of anxiety, respectively). Manipulating the contingency of threat cues is thought to implicitly train the individual to attend away from threat cues or toward safety cues. For example, Amir and colleagues (Amir, Beard, Burns, & Bomyea, 2009; Amir, Beard, Taylor, et al., 2009) randomized individuals diagnosed with generalized anxiety disorder into either ABM or a control condition. After eight sessions, the ABM group showed significant reductions in attentional bias toward threat and anxiety, as evaluated by self-reports and clinical interview (Amir et al., 2009).

The two sets of studies outlined, correlational and experimental, help the field create a more complex understanding of the potential relation between attention to threat and anxiety, providing evidence for multiple forms of validity. This approach also can point to concerns that would not otherwise be evident. That is, although some meta-analyses of ABM studies provided support for the causal role of attention

(Bar-Haim, 2010; Hakamata et al., 2010), others questioned the reliability and breadth of the relation by finding important moderators (e.g., ABM delivered in the clinic versus at home; Cristea, Kok, & Cuijpers, 2015; Heeren, Mogoaşe, Philippot, & McNally, 2015; Linetzky, Pergamin-Hight, Pine, & Bar-Haim, 2015; Mogoaşe, David, & Koster, 2014). Questions are also evident from more correlational data. In the only longitudinal study to our knowledge, attention bias to threat at age 5 failed to predict later anxiety at age 7 (White et al., 2017). Rather, concurrent affect-biased attention toward both threat and reward moderated the relation between early fearful temperament and anxiety, such that early fearful temperament predicted anxiety only for children who displayed a bias toward threat or those who did not display a bias toward reward.

Luckily, we can bring together experimental and individual difference (person-centered) approaches to probe these emerging questions further. For example, Morales, Taber-Thomas, and Pérez-Edgar (2017) had children complete both an affective Posner variant and a dot-probe task. They found no significant correlations in attention to threat across tasks for the full sample. However, behaviorally inhibited children did show a cross-task correlation in attention to threat. Morales et al. then classified the children as showing, or not showing, a stable pattern across tasks (either threat vigilance or threat avoidance). As expected, the stable group was dominated by children high in behavioral inhibition. Importantly, it was within the stable groups that children showed elevated levels of anxiety. Thus, Morales et al. brought together natural individual variation, experimental manipulation, and variable- and person-centered approaches to note subtle patterns in questions of interest and to address open issues in the larger literature. This relatively small, cross-sectional study can serve as the foundation for larger, long-term systematic studies.

The fits and starts of this literature, with novel findings that later seem weaker (pessimistic view) or more nuanced (optimistic view) than initially thought, reflect broader trends. That is, these data are situated within a broader call in the field for a more robust and reproducible science (Lindsay, 2015; Open Science Collaboration, 2015). We would argue that the issues raised in this chapter speak to these concerns. This review cannot touch on all important issues, such as publication trends for nonsignificant findings and questionable statistical methods in search of significance. However, we suggest that the use of multiple methods in and out of the laboratory can allow us to probe a core shared question from multiple points of view. Including an individual differences approach then allows us to see when and how and for whom our "settled" answer actually applies. Just as individual puzzle pieces may look quite different, it is unrealistic to believe that every probe will produce the same answer. However, when brought together, the puzzle pieces should form a coherent, and three-dimensional, portrait of both emotions (our constructs of interest) and the child in which they reside.

Our field's current difficulty in reproducibility may be in small part a reflection of our overreliance on "average" and "normative." From this point of view, experimental manipulations always should produce the same outcome, across individuals and across time. Individual differences are thought to create bothersome noise that should be titrated out through clever experimental design. However, when replication fails, the specter of the "hidden" moderator is raised. We suggest that this approach may be incorrect both at outset and in the post mortem. Our review suggests that the strength of our science, regardless of the current label applied, lies in its ability to cross theoretical and empirical divides to incorporate multiple methods and the worldviews that fuel them. In this way we can better capture our elusive target—the ever-evolving emotional child moving through space and time, working to shape the world just as the world changes the child.

REFERENCES

Amaral, D., Price, J., Pitkanen, A., & Carmichael, S. (1992). Anatomical organization of the primate amygdaloid complex. In *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction* (pp. 1–66). New York, NY: Wiley.

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology*, *109*, 28–33. doi:10.137/a0012589

Amir, N., Beard, C., Taylor, C. T. K., H., Elias, J., Bums, M., & Chen, X. (2009). Attention training in individuals with generalized social phobia: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*, 961–973. doi:10.1037/a0016685

Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: the case of laboratory aggression. *Review of General Psychology*, *1*, 19.

Andison, F. S. (1977). TV violence and viewer aggression: a cumulation of study results 1956–1976. *Public Opinion Quarterly*, *41*, 314–331.

Aslin, R. N., & Fiser, J. (2005). Methodological challenges for understanding cognitive development in infants. *Trends in Cognitive Sciences*, *9*, 92–98.

Bar-Haim, Y. (2010). Research review: attention bias modification (ABM): a novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry*, *51*(8), 859–870.

Bar-Haim, Y., Fox, N. A., Benson, B., Guyer, A. E., Williams, A., Nelson, E. E., . . . Ernst, M. (2009). Neural correlates of reward processing in adolescents with a history inhibited temperament. *Psychological Science*, *20*, 1009–1018. doi:10.1111/j.1467–9280.2009.02401.x

Bar-Haim, Y., Fox, N. A., VanMeenen, K. M., & Marshall, P. J. (2004). Children's narratives and patterns of cardiac reactivity. *Developmental Psychobiology*, *44*, 238–249. doi:10.1002/dev.20006

Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition & Emotion*, *12*, 579–599.

Bell, M. A., & Cuevas, K. (2012). Using EEG to study cognitive development: issues and practices. *Journal of Cognition and Development*, *13*(3), 281–294.

Bell, M. A., & Wolfe, C. D. (2007). The cognitive neuroscience of early socioemotional development. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional Development in the Toddler Years: Transitions and Transformations* (pp. 345–369). New York, NY: Guilford Press.

Bjorklund, D. F. (1997). In search of a metatheory for cognitive development (or, Piaget is dead and I don't feel so good myself). *Child Development*, *68*, 144–148.

Blackford, J. U., & Pine, D. S. (2012). Neural substrates of childhood anxiety disorders: a review of neuroimaging findings. *Child and Adolescent Psychiatric Clinics of North America*, *21*, 501–525. doi:10.1016/j.chc.2012.05.002

Boring, E. G. (1929). *History of Experimental Psychology*. N.p.: Genesis Publishing.

Brody, L. R. (1985). Gender differences in emotional development: a review of theories and research. *Journal of Personality*, *53*, 102–149.

Buss, K. A., & Goldsmith, H. H. (2000). *Manual and Normative Data for the Laboratory Temperament Assessment Battery—Toddler Version*. Department of Psychology Technical Report, University of Wisconsin.

Calkins, S. D., Fox, N. A., & Marshall, T. R. (1996). Behavioral and physiological antecedents of inhibited and uninhibited behavior. *Child Development*, *67*, 523–540.

Casey, B. J., Trainor, R. J., Orendi, J. L., Schubert, A. B., Nystrom, L. E., Giedd, J. N., . . . Rapoport, J. L. (1997). A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. *Journal of Cognitive Neuroscience*, *9*, 835–847. doi:10.1162/jocn.1997.9.6.835

Cattell, J. M. (1898). The biological problems of today: psychology. *Science*, *7*, 152–154.

Chantala, K., & Tabor, J. (2010). *National Longitudinal Study of Adolescent Health: Strategies to Perform a Design-based Analysis Using the Add Health Data*. Retrieved from University of North Carolina at Chapel Hill. http://citeseerx.ist.psu.edu

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: a meta-analytic review. *Psychological Bulletin*, *139*, 735–765.

Chaplin, T. M., Cole, P. M., & Zahn-Waxler, C. (2005). Parental socialization of emotion expression: gender differences and relations to child adjustment. *Emotion*, *5*, 80.

Chronis-Tuscano, A., Degnan, K. A., Pine, D. S., Pérez-Edgar, K., Henderson, H. A., Diaz, Y., . . . Fox, N. A. (2009). Stable behavioral inhibition during infancy and early childhood predicts the development of anxiety disorders in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*, 928–935. doi:10.1097/CHI.ob013e3181ae09df

Clauss, J. A., & Blackford, J. U. (2012). Behavioral inhibition and risk for developing social anxiety disorder: a meta-analytic study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*, 1066–1075. doi:10.106/j.jaac

Cohen, G. L., & Prinstein, M. J. (2006). Peer contagion of aggression and health risk behavior among adolescent males: an experimental investigation of effects on public conduct and private attitudes. *Child Development*, *77*, 967–983.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology* (5th ed.), *Social, Emotional, and Personality Development* (Vol. 3, pp. 779–862). New York, NY: Wiley.

Cole, C., Zapp, D. J., Fettig, N. B., & Pérez-Edgar, K. (2016). Impact of attention biases to threat and effortful control on individual variations in negative affect and social withdrawal in very young children. *Journal of Experimental Child Psychology*, *141*, 210–221.

Cole, P. M. (1986). Children's spontaneous control of facial expression. *Child Development*, *57*, 1309–1321.

Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: methodological challenges and directions for child development research. *Child Development*, *75*, 317–333.

Coplan, R. J., Rubin, K. H., Fox, N. A., Calkins, S. D., & Stewart, S. L. (1994). Being alone, playing alone, and acting alone: distinguishing among reticence and passive and active solitude in young children. *Child Development*, *65*, 129–137. doi:10.1111/j.1467-8624.1994.tb00739.x

Crick, N. R., Nelson, D. A., Morales, J. R., Cullerton-Sen, C., Casas, J. F., & Hickman, S. E. (2001). Relational victimization in childhood and adolescence. In J. Juvonen & S. Graham (Eds.), *Peer Harassment in School: The Plight of the Vulnerable and Victimized* (pp. 196–214). New York, NY: Guilford Press.

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *British Journal of Psychiatry 206*(1), 7−16.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671.

Cuevas, K., & Bell, M. A. (2011). EEG and ECG from 5 to 10 months of age: developmental changes in baseline activation and cognitive processing during a working memory task. *International Journal of Psychophysiology*, *80*(2), 119–128.

Damon, W., Kuhn, D., & Siegler, R. S. (Eds.). (1998). *Handbook of child psychology* (5th ed., Vol. 2), *Cognition, Perception and Language*. Chichester, UK: Wiley.

Davis, M., Walker, D. L., & Lee, Y. (1997). Amygdala and bed nucleus of the stria terminalis: differential roles in fear and anxiety measured with the acoustic startle reflex. *Philosophical*

Transactions of the Royal Society of London Series B, Biological Sciences, *352*, 1675–1687.

De Los Reyes, A. (2011). More than measurement error: discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *40*, 1–9.

Denham, S. A. (1998). *Emotional Development in Young Children*. New York, NY: Guilford Press.

Dennis, T. A. (2006). Emotional self-regulation in preschoolers: the interplay of child approach reactivity, parenting, and control capacities. *Developmental Psychology*, *42*, 84.

Dennis, T. A., Buss, K. A., & Hastings, P. D. (2012). Introduction to the monograph: physiological measures of emotion from a developmental perspective: state of the science. *Monographs of the Society for Research in Child Development*, *77*, 1–5.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, *111*, 225–236.

Deveney, C. M., Connolly, M. E., Haring, C. T., Bones, B. L., Reynolds, R. C., Kim, P., . . . Leibenluft, E. (2013). Neural mechanisms of frustration in chronically irritable children. *JAMA Psychiatry*, *170*, 1186–1194.

Dodge, K. A. (1991). Emotion and social information processing. In J. Garber & K. Dodge (Eds.), *The Development of Emotion Regulation and Dysregulation* (pp. 159–181). New York, NY: Cambridge University Press.

Durston, S., Thomas, K. M., Yang, Y., Ulug, A. M., Zimmerman, R. D., & Casey, B. J. (2002). A neural basis for the development of inhibitory control. *Developmental Science*, *5*(4), F9–F16.

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, *51*, 665–697.

Eisenberg, N., Champion, C., & Ma, Y. (2004). Emotion-related regulation: an emerging construct. *Merrill-Palmer Quarterly*, *50*, 236–259. doi:10.1056/NEJMoa0908629

Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: sharpening the definition. *Child Development*, *75*, 334–339. doi:10.1111/j.1467-8624.2004.00674.x

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. J. Power (Eds.). *Handbook of Cognition and Emotion* (pp. 45–60). Chichester, UK: Wiley.

Eldar, S., Ricon, T., & Bar-Haim, Y. (2008). Plasticity in attention: implications for stress response in children. *Behaviour Research and Therapy*, *46*(4), 450–461.

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: a meta-analysis. *Psychological Bulletin*, *132*(1), 33.

Eriksen, C. (1995). The flankers task and response competition: a useful tool for investigating a variety of cognitive problems. In C. Bundesen & H. Shibuya (Eds.), *Visual Selective Attention* (pp. 101–118). Hillsdale, NJ: Erlbaum.

Forbes, E. E., Fox, N. A., Cohn, J. F., Galles, S. F., & Kovacs, M. (2006). Children's affect regulation during a disappointment: psychophysiological responses and relation to parent history of depression. *Biological Psychology*, *71*(3), 264–277.

Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental Psychology*, *46*, 1454.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, *130*(4), 681–700.

Fox, N. A., Hane, A. A., & Pérez-Edgar, K. (2006). Psychophysiological methods for the study of developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental Psychopathology* (Vol. 2, pp. 381–426). Hoboken, NJ: Wiley.

Fox, N. A., & Helfinstein, S. M. (2013). The contribution of temperament to the study of social cognition: learning whether the glass is half empty or half full. In M. R. Banaji & S. Gelman (Eds.), *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us* (pp. 49–53). Oxford, UK: Oxford University Press.

Fox, N. A., Kirwan, M., & Reeb-Sutherland, B. (2012). Measuring the physiology of emotion and emotion regulation—timing is everything. *Monographs of the Society for Research in Child Development*, *77*, 98–108.

Fox, N. A., Rubin, K. H., Calkins, S. D., Marshall, T. R., Coplan, R. J., Porges, S. W., . . . Stewart, S. (1995). Frontal activation asymmetry and social competence at four years of age. *Child Development*, *66*, 1771–1784. doi:10.1111/j.1467-8624.1995.tb00964.x

Fox, N. A., Snidman, N., Haas, S. A., Degnan, K. A., & Kagan, J. (2015). The relations between reactivity at 4 months and behavioral inhibition in the second year: replication across three independent samples. *Infancy*, *20*(1), 98–114.

Fry, D. P., & Gabriel, A. H. (1994). Preface: the cultural construction of gender and aggression. *Sex Roles*, *30*, 165–167.

Garcia Coll, C., Kagan, J., & Reznick, J. S. (1984). Behavioral inhibition in young children. *Child Development*, *55*, 1005–1019. doi:10.2307/1130152

Garrett-Peters, P. T., & Fox, N. A. (2007). Cross-cultural differences in children's emotional reactions to a disappointing situation. *International Journal of Behavioral Development*, *31*, 161–169.

Gee, D. G., Humphreys, K. L., Flannery, J., Goff, B., Telzer, E. H., Shapiro, M., . . . Tottenham, N. (2013). A developmental shift from positive to negative connectivity in human amygdala–prefrontal circuitry. *Journal of Neuroscience*, *33*, 4584–4593.

Gerardi-Caulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24–36 months of age. *Developmental Science*, *3*, 397–404. doi:10.1111/1467-7687.00134

Gerstadt, C. L., Hong, Y. H., & Diamond, A. (1994). The relationship between cognition and action: performance of children 3 ½–7 years old on a Stroop-like day-night test. *Cognition*, *53*, 129–153.

Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F., Liu, H., & Zijdenbos, A. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, *2*, 861–863.

Grossmann, T. (2015). The development of social brain functions during infancy. *Psychological Bulletin*, *141*, 1266–1287.

Guyer, A. E., Benson, B., Choate, V. R., Bar-Haim, Y., Pérez-Edgar, K., Jarcho, J. M., . . . Nelson, E. E. (2014). Lasting associations between early-childhood temperament and late-adolescent reward-circuitry response to peer feedback. *Development and Psychopathology*, *26*, 229–243.

Guyer, A. E., Jarcho, J. M., Pérez-Edgar, K., Degnan, K. A., Pine, D. S., Fox, N. A., & Nelson, E. E. (2015). Temperament and parenting styles in early childhood differentially influence neural response to peer evaluation in adolescence. *Journal of Abnormal Child Psychology*, *43*, 863–874.

Guyer, A. E., Nelson, E. E., Pérez-Edgar, K., Hardin, M. G., Roberson-Nay, R., Monk, C. S., . . . Ernst, M. (2006). Striatal functional alteration in adolescents characterized by early childhood behavioral inhibition. *Journal of Neuroscience*, *26*(24), 6399–6405.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., . . . Pine, D. S. (2010). Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biologival Psychiatry*, *68*(11), 982–990.

Han, G., Miller, J. G., Cole, P. M., Zahn-Waxler, C., & Hastings, P. D. (2015). Adolescents' internalizing and externalizing problems predict their affect-specific HPA and HPG axes reactivity. *Developmental Psychobiology*, *57*, 769–785.

Hane, A. A., Fox, N. A., Henderson, H. A., & Marshall, P. J. (2008). Behavioral reactivity and approach-withdrawal bias in infancy. *Developmental Psychology*, *44*, 1491–1496.

Hare, T. A., Tottenham, N., Galvan, A., Voss, H. U., Glover, G. H., & Casey, B. J. (2008). Biological substrates of emotional reactivity and regulation in adolescence during an emotional go-no go task. *Biological Psychiatry*, *63*, 927–934.

Harlow, H. F. (1958). The nature of love. *American Psychologist*,*13*, 573–685.

Harris, P. L. (1989). *Children and Emotion: The Development of Psychological Understanding*. Cambridge, MA: Basil Blackwell.

Hastings, P. D., Kahle, S., & Nuselovici, J. M. (2014). How well socially wary preschoolers fare over time depends on their parasympathetic regulation and socialization. *Child Development*, *85*, 1586–1600.

Hastings, P. D., Kahle, S. S., & Han, G. H. P. (2014). Developmental affective psychophysiology: using physiology to inform our understanding of emotional development. In K. Lagattuta (Ed.), *Children and Emotion: New Insights into Developmental Affective Science* (Vol. 26, pp. 13–28). Basel, Switzerland: Karger.

Hastings, P. D., Miller, J. G., Kahle, S., & Zahn-Waxler, C. (2014). The neurobiological bases of empathic concern for others. In M. Killen & J. Smetana (Eds.), *Handbook of Moral Development* (pp. 411–434). New York, NY: Psychology Press.

Heeren, A., Mogoaşe, C., Philippot, P., & McNally, R. J. (2015). Attention bias modification for social anxiety: a systematic review and meta-analysis. *Clinical Psychology Review*, *40*, 76–90.

Helfinstein, S. M., Benson, B., Pérez-Edgar, K., Bar-Haim, Y., Detloff, A., Pine, D. S., . . . Ernst, M. (2011). Striatal responses to negative monetary outcomes differ between behaviorally inhibited and non-inhibited adolescents. *Neuropsychologia*, *49*, 479–485. doi:10.1016/j.neuropsychologia.2010.12.015

Hoehl, S., & Striano, T. (2010). The development of emotional face and eye gaze processing. *Developmental Science*, *13*(6), 813–825.

Hofer, C., Eisenberg, N., Spinrad, T. L., Morris, A. S., Gershoff, E., Valiente, C., . . . Eggum, N. D. (2013). Mother–adolescent conflict: stability, change, and relations with externalizing and internalizing behavior problems. *Social Development*, *22*, 259–279.

Izard, C. E. (1993). Organizational and motivational functions of discrete emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of Emotions* (pp. 631–641). New York, NY: Guilford Press.

Jouriles, E. N., Murphy, C. M., & O’Leary, K. D. (1989). Effects of maternal mood on mother-son interaction patterns. *Journal of Abnormal Child Psychology*, *17*, 513–525.

Kagan, J. (1994). On the nature of emotion. *Monographs of the Society for Research in Child Development*, *59*(2–3), 7–24, 250–283.

Kagan, J. (2012). The biography of behavioral inhibition. In M. Zentner & R. L. Shiner (Eds.), *Handbook of Temperament* (pp. 69–82). New York, NY: Guilford Press.

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Development*, *58*, 1459–1473. doi:10.2307/1130685

Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science*, *240*, 167–171.

Kagan, J., Snidman, N., McManis, M., & Woodward, S. (2001). Temperamental contributions to the affect family of anxiety. *Psychiatric Clinics of North America*, *24*(4), 677–688. doi:10.1016/S0193-953X(05)70257–4

Kahle, S., Miller, J. G., Lopez, M., & Hastings, P. D. (2016). Sympathetic recovery from anger is associated with emotion regulation. *Journal of Experimental Child Psychology*, *142*, 359–371.

Khoo, S. T., West, S. G., Wu, W., & Kwok, O. M. (2006). Longitudinal methods. In M. Eid & E. Diener (Eds.), *Handbook of Multimethod Measurement in Psychology* (pp. 301–317). Washington, DC: American Psychological Association.

Kieras, J. E., Tobin, R. M., Graziano, W. G., & Rothbart, M. K. (2005). You can’t always get what you want: effortful control and children’s responses to undesirable gifts. *Psychological Science*, *16*, 391–396.

Klein, D. N., Dyson, M. W., Kujawa, A., & Kotov, R. (2012). Temperament and internalizing disorders. In M. Zentner & R. L. Shiner (Eds.), *Handbook of Temperament* (pp. 541–561). New York, NY: Guilford Press.

Klimes-Dougan, B., Brand, A. E., Zahn-Waxler, C., Usher, B., Hastings, P. D., Kendziora, K., & Garside, R. B. (2007). Parental emotion socialization in adolescence: differences in sex, age and problem status. *Social Development*, *16*, 326–342.

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: expression, experience, and physiology. *Journal of Personality and Social Psychology*, *74*, 686.

Lagattuta, K. H., Sayfan, L., & Monsour, M. (2011). A new measure for assessing executive function across a wide age range: children and adults find happy-sad more difficult than day-night. *Developmental Science*, *14*, 481–489.

Lahat, A., Pérez-Edgar, K., Degnan, K. A., Guyer, A. E., Lejuez, C. W., Ernst, M., . . . Fox, N. A. (2012). Early childhood temperament predicts substance use in young adults. *Translational Psychiatry*, *2*, e157.

Larsen, R. J., & Prizmic-Larsen, Z. (2006). Measuring emotions: implications of a multimethod perspective. In M. Eid & E. Diener (Eds.), *Handbook of Multimethod Measurement in Psychology* (pp. 337–351). Washington, DC: American Psychological Association.

Leibenluft, E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *American Journal of Psychiatry*, *168*, 129–142.

Leibenluft, E., & Rich, B. A. (2008). Pediatric bipolar disorder. *Annual Review of Clinical Psychology*, *4*, 163–187.

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, *10*(1), 37–47.

Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created "social climates." *Journal of Social Psychology*, *10*, 269–299.

Lewis, M. (2010). The emergence of human emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (pp. 304–319). New York, NY: Guilford Press.

Lewis, M. D., Hitchcock, D. F. A., & Sullivan, M. W. (2004). Physiological and emotional reactivity to learning and frustration. *Infancy*, *6*, 121–143.

Lewis, M. D., Lamm, C., Segalowitz, S. J., Stieben, J., & Zelazo, P. D. (2006). Neurophysiological correlates of emotion regulation in children and adolescents. *Journal of Cognitive Neuroscience*, *18*, 430–443.

Lewis, M. D., & Stieben, J. (2004). Emotion regulation in the brain: conceptual issues and directions for developmental research. *Child Development*, *75*, 371–376.

Lindsay, D. S. (2015). Replication in psychological science. *Psychological Science*, *26*, 1827–1832.

Linetzky, M., Pergamin-Hight, L., Pine, D. S., & Bar-Haim, Y. (2015). Quantitative evaluation of the clinical efficacy of attention bias modification treatment for anxiety disorders. *Depression and Anxiety*, *32*(6), 383–391.

LoBue, V., & Pérez-Edgar, K. (2014). Sensitivity to social and non-social threats in temperamentally shy children at-risk for anxiety. *Developmental Science*, *17*, 239–247.

LoBue, V., Rakison, D., & DeLoache, J. S. (2010). Threat perception across the lifespan: evidence for multiple converging pathways. *Current Directions in Psychological Science*, *19*, 375–379.

Loeber, R., & Farrington, D. P. (1994). Problems and solutions in longitudinal and experimental treatment studies of child psychopathology and delinquency. *Journal of Consulting and Clinical Psychology*, *62*, 887.

Luna, B., & Sweeney, J. A. (2004). The emergence of collaborative brain function: fMRI studies of the development of response inhibition. *Annals of the New York Academy of Sciences*, *1021*, 296–309. doi:10.1196/annals.1308.035

Maccoby, E. E. (1998). *The Two Sexes: Growing Up Apart, Coming Together*. Cambridge, MA: Harvard University Press.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*, 15–20. doi:10.1037//0021-843X.95.1.15

Marceau, K., Shirtcliff, E. A., Hastings, P. D., Klimes-Dougan, B., Zahn-Waxler, C., Dorn, L. D., & Susman, E. J. (2014). Within-adolescent coupled changes in cortisol with DHEA and testosterone in response to three stressors during adolescence. *Psychoneuroendocrinology*, *41*, 33–45.

Marshall, P. J., Reeb, B. C., & Fox, N. A. (2009). Electrophysiological responses to auditory novelty in temperamentally different 9-month-old infants. *Developmental Science*, *12*, 568–582.

Marshall, P. J., & Stevenson-Hinde, J. (1998). Behavioral inhibition, heart period, and respiratory sinus arrhythmia in young children. *Developmental Psychobiology*, *33*(3), 283–292.

Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, *23*(5), 563–569.

McCall, R. B. (1977). Challenges to a science of developmental psychology. *Child Development*, *48*(2), 333–344.

McClure, E. B., Monk, C. S., Nelson, E. E., Parrish, J. M., Adler, A., Blair, R. J. R., … Pine, D. S. (2007). Abnormal attention modulation of fear circuit function in pediatric generalized anxiety disorder. *Archives of General Psychiatry*, *64*, 109–116.

McDermott, J. M., Pérez-Edgar, K., Henderson, H. A., Chronis-Tuscano, A., Pine, D. S., & Fox, N. A. (2009). A history of childhood behavioral inhibition and enhanced response monitoring in adolescence are linked to clinical anxiety. *Biological Psychiatry*, *65*, 445–448. doi:10.1016/j.biopsych.2008.10.043

McDowell, D. J., Kim, M., O'Neil, R., & Parke, R. D. (2002). Children's emotional regulation and social competence in middle childhood: the role of maternal and paternal interactive style. *Marriage & Family Review*, *34*, 345–364.

Measelle, J. R., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (1998). Assessing young children's views of their academic, social, and emotional lives: an evaluation of the Self-Perception Scales of the Berkeley Puppet Interview. *Child Development*, *69*, 1556–1576.

Miskovic, V., & Schmidt, L. A. (2012). Early information processing biases in social anxiety. *Cognition & Emotion*, *26*, 176–185.

Mogoaşe, C., David, D., & Koster, E. H. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology*, *70*(12), 1133–1157.

Morales, S., Fu, X., & Pérez-Edgar, K. (2016). A developmental neuroscience perspective on affect biased attention. *Developmental Cognitive Neuroscience*, *21*, 26–41.

Morales, S., Pérez-Edgar, K., & Buss, K. A. (2015). Attention biases towards and away from threat mark the relation between early dysregulated fear and the later emergence of social withdrawal. *Journal of Abnormal Child Psychology*, *43*, 1067–1078.

Morales, S., Taber-Thomas, B. C., & Pérez-Edgar, K. (2017). Patterns of attention to threat across tasks in behaviorally inhibited children at risk for anxiety. *Developmental Science*, *20*, e12391.

Morris, A. S., Robinson, L. R., & Eisenberg, N. (2006). Applying a multimethod perspective to the study of developmental psychology. In M. Eid & E. Diener (Eds.), *Handbook of Multimethod Measurement in Psychology* (pp. 371–384). Washington, DC: American Psychological Association.

Nesse, R. M., Curtis, G. C., Thyer, B. A., McCann, D. S., Huber-Smith, M. J., & Knopf, R. F. (1985). Endocrine and cardiovascular responses during phobic anxiety. *Psychosomatic Medicine*, *47*, 320–332.

Newman, J. (1985). The infant cry of primates: an evolutionary perspective. In C. F. Z. Boukydis & B. M. Lester (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 307–323). New York, NY: Plenum Press.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. doi:10.1126/science.aac4716

Overton, W. F. (2006). Developmental psychology: philosophy, concepts, methodology. In R. M. Lerner & W. Damon (Eds.), *Handbook of Child Psychology* (pp. 18–88). Hoboken, NJ: Wiley.

Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental science*, *11*(1), 60–68.

Pepler, D. J., & Craig, W. M. (1995). A peek behind the fence: naturalistic observations of aggressive children with remote audiovisual recording. *Developmental Psychology*, *31*, 548.

Pérez-Edgar, K. (2015). Effortful control in adolescence: individual differences within a unique developmental window. In G. Oettingen & P. M. Gollwitzer (Eds.), *Self-regulation in Adolescence* (pp. 78–102). Cambridge UK: Cambridge University Press.

Pérez-Edgar, K., & Bar-Haim, Y. (2010). Application of cognitive-neuroscience techniques to the study of anxiety-related processing biases in children. In J. Hadwin & A. Field (Eds.), *Information Processing Biases in Child and Adolescent Anxiety* (pp. 183–206). Chichester, UK: Wiley.

Pérez-Edgar, K., Bar-Haim, Y., McDermott, J. M., Chronis-Tuscano, A., Pine, D. S., & Fox, N. A. (2010). Attention biases to threat and behavioral inhibition in early childhood shape adolescent social withdrawal. *Emotion*, *10*, 349–357. doi:10.1037/a0018486

Pérez-Edgar, K., & Fox, N. A. (2003). Individual differences in children's performance during an emotional Stroop task: a behavioral and electrophysiological study. *Brain and Cognition*, *52*, 33–51. doi:10.1016/S0278-2626(03)00007-1

Pérez-Edgar, K., & Fox, N. A. (2005). A behavioral and electrophysiological study of children's selective attention under neutral and affective conditions. *Journal of Cognition and Development*, *6*(1), 89–118.

Pérez-Edgar, K., Fox, N. A., Cohn, J. F., & Kovacs, M. (2006). Behavioral and electrophysiological markers of selective attention in children of parents with a history of depression. *Biological Psychiatry*, *60*, 1131–1138.

Pérez-Edgar, K., Hardee, J., Guyer, A. E., Benson, B., Nelson, E. E., Gorodetsky, E., . . . Ernst, M. (2014). DRD4 and striatal modulation of the link between childhood behavioral inhibition and adolescent anxiety. *Social, Cognitive, and Affective Neuroscience*, *9*, 445–453. doi:10.1093/scan/nst001

Pérez-Edgar, K., Reeb-Sutherland, B. C., McDermott, J. M., White, L. K., Henderson, H. A., Degnan, K. A., . . . Fox, N. A. (2011). Attention biases to threat link behavioral inhibition to social withdrawal over time in very young children. *Journal of Abnormal Child Psychology*, *39*, 885–895. doi:10.1007/s10802-011-9495-5

Pérez-Edgar, K., Roberson-Nay, R., Hardin, M. G., Poeth, K., Guyer, A. E., Nelson, E. E., . . . Ernst, M. (2007). Attention alters neural responses to evocative faces in behaviorally inhibited adolescents. *NeuroImage*, *35*, 1538–1546.

Pérez-Edgar, K., Schmidt, L. A., Henderson, H. A., Schulkin, J., & Fox, N. A. (2008). Salivary cortisol levels and infant temperament shape developmental trajectories in boys at risk for behavioral maladjustment. *Psychoneuroendocrinology*, *33*, 916–925.

Pérez-Edgar, K., Taber-Thomas, B., Auday, E., & Morales, S. (2014). Temperament and attention as core mechanisms in the early emergence of anxiety. In K. Lagattuta (Ed.), *Children and Emotion: New Insights into Developmental Affective Science* (Vol. 26, pp. 42–56). Basel, Switzerland: Karger.

Petrill, S. A., & Brody, N. (2002). Personality and individual differences. In H. Pashler (Ed.), *Stevens' Handbook of Experimental Psychology*, (Vol. 4, pp. 563–600). Hoboken, NJ: Wiley.

Polivy, J., & Doyle, C. (1980). Laboratory induction of mood states through the reading of self-referent mood statements: affective changes or demand characteristics? *Journal of Abnormal Psychology*, *89*, 286.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Boushuis (Eds.), *Attention and Performance X* (pp. 531–556), London, UK: Erlbaum.

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of Experimental Child Psychology*, *108*, 621–637. doi:10.1016/j.jecp.2010.09.008

Reeb-Sutherland, B. C., Helfinstein, S. M., Degnan, K. A., Pérez-Edgar, K., Henderson, H. A., Lissek, S., . . . Fox, N. A. (2009). Startle modulation in behaviorally inhibited adolescents with a lifetime occurrence of anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*, 610–617.

Reeb-Sutherland, B. C., Vanderwert, R. E., Degnan, K. A., Marshall, P. J., Pérez-Edgar, K., Chronis-Tuscano, A., . . . Fox, N. A. (2009). Attention to novelty in behaviorally inhibited adolescents moderates risk for anxiety. *Journal of Child Psychology and Psychiatry*, *50*, 1365–1372.

Reeb-Sutherland, B. C., Williams, L. R., Degnan, K. A., Pérez-Edgar, K., Chronis-Tuscano, A., Leibenluft, E., . . . Fox, N. A. (2015). Identification of emotional facial expressions among behaviorally inhibited adolescents with lifetime anxiety disorders. *Cognition & Emotion*, *29*, 372–382.

Reese, H. W., & Lipsitt, L. P. (1970). *Experimental Child Psychology*. New York, NY: Academic Press.

Rich, B. A., Schmajuk, M., Pérez-Edgar, K., Fox, N. A., Pine, D. S., & Leibenluft, E. (2007). Different psychophysiological and behavioral responses elicited by frustration in pediatric bipolar disorder and severe mood dysregulation. *American Journal of Psychiatry*, *164*(2), 309–317.

Rich, B. A., Schmajuk, M., Pérez-Edgar, K., Pine, D. S., Fox, N. A., & Leibenluft, E. (2005). The impact of reward, punishment, and frustration on attention in pediatric bipolar disorder. *Biological Psychiatry*, *58*, 532–539.

Roy, A. K., Dennis, T. A., & Warner, C. M. (2015). A critical review of attentional threat bias and its role in the treatment of pediatric anxiety disorders. *Journal of Cognitive Psychotherapy*, *29*(3), 171–184.

Rubin, K., Chen, X., & Hymel, S. (1993). Socioemotional characteristics of withdrawn and aggressive children. *Merrill-Palmer Quarterly*, *39*, 518–534.

Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child Development*, *55*(4), 1504–1513.

Scarr, S. (1992). Developmental theories for the 1990s: development and individual differences. *Child Development*, *63*, 1–19.

Scarr, S., & McCartney, K. (1983). How people make their own environments: a theory of genotype-environmental effects. *Child Development*, *54*, 424–435.

Schmidt, L. A., Fox, N. A., Rubin, K. H., Sternberg, E. M., Gold, P. W., Smith, C. C., & Schulkin, J. (1997). Behavioral and neuroendocrine responses in shy children. *Developmental Psychobiology*, *30*, 127–140.

Schmidt, L. A., Fox, N. A., Schulkin, J., & Gold, P. W. (1999). Behavioral and psychophysiological correlates of self-presentation in temperamentally shy children. *Developmental Psychobiology*, *35*, 119–135. doi:10.1002/(SICI)1098-2302(199909)35:2<119::AID-DEV5>3.0.CO;2-G

Schmidt, L. A., Fox, N. A., Sternberg, E. M., Gold, P. W., Smith, C. C., & Schulkin, J. (1999). Adrenocortical reactivity and social competence in seven-year-olds. *Personality and Individual Differences*, *26*(6), 977–985.

Schmidt, L. A., & Segalowitz, S. J. (2008). *Developmental Psychophysiology: Theory, Systems, and Methods*. New York, NY: Cambridge University Press.

Schwartz, C. E., Wright, C. I., Shin, L. M., Kagan, J., & Rauch, S. L. (2003). Inhibited and uninhibited infants "grown up": adult amygdalar response to novelty. *Science*, *300*, 1952–1953.

Shechner, T., Britton, J. C., Pérez-Edgar, K., Bar-Haim, Y., Ernst, M., Fox, N. A., . . . Pine, D. S. (2012). Attention biases, anxiety, and development: toward or away from threats or rewards? *Depression & Anxiety*, *29*, 282–294. doi:10.1002/da.20914

Smith, R. H., & Harris, M. J. (2006). Multimethod approaches in social psychology: between- and within-method replication and multimethod assessment. In M. Eid & E. Diener (Eds.), *Handbook of Multimethod Measurement in Psychology* (pp. 385–400). Washington, DC: American Psychological Association.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. doi:10.1037/h0054651

Talwar, V., Murphy, S. M., & Lee, K. (2007). White lie-telling in children for politeness purposes. *International Journal of Behavioral Development*, *31*, 1–11.

Thompson, R. A. (2011). Emotion and emotion regulation: two sides of the developing coin. *Emotion Review*, *3*, 53–61.

Tobin, R. M., & Graziano, W. G. (2011). The disappointing gift: dispositional and situational moderators of emotional expressions. *Journal of Experimental Child Psychology*, *110*, 227–240.

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012). Affect-biased attention as emotion regulation. *Trends in Cognitive Sciences*, *16*, 365–372. doi:10.1016/j.tics.2012.06.003

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist*, *44*, 112.

Underwood, M. K., Galen, B. R., & Paquette, J. A. (2001). Top ten challenges for understanding gender and aggression in children: why can't we all just get along? *Social Development*, *10*, 248–266.

van der Molen, M. W., & Molenaar, P. C. M. (1994). Cognitive psychophysiology: a window to cognitive development and brain maturation. In G. Dawson & K. W. Fischer (Eds.), *Human Behavior and the Developing Brain* (pp. 456–491). New York, NY: Guilford Press.

Wallace, D. B., Franklin, M. B., & Keegan, R. T. (1994). The observing eye: a century of baby diaries. *Human Development*, *37*, 1–29.

Waters, A. M., Mogg, K., Bradley, B. P., & Pine, D. S. (2008). Attentional bias for emotional faces in children with generalized anxiety disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(4), 435–442.

Watson, J. B. (1926). Behaviourism: a psychology based on reflex-action. *Philosophy*, *1*, 454–466.

White, L. K., Degnan, K. A., Henderson, H. A., Pérez-Edgar, K., Walker, O. L., Schechner, T., . . . Fox, N. A. (2017). Developmental relations between behavioral inhibition, anxiety, and attention biases to threat and positive information. *Child Development*, *88*, 141–155.

White, L. K., Lamm, C., Helfinstein, S. M., & Fox, N. A. (2012). Neurobiology and neurochemistry of temperament in children. In M. Zentner & R. L. Shiner (Eds.), *Handbook of Temperament* (pp. 347–367). New York, NY: Guilford Press.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: statistical models and methodological recommendations. *Development and Psychopathology*, *10*, 395–426.

Wilson, E. J., & MacLeod, C. (2003). Contrasting two accounts of anxiety-linked attentional bias: selective attention to varying levels of stimulus threat intensity. *Journal of Abnormal Psychology*, *112*(2), 212–218.

Wohlwill, J. F. (1973). *The Study of Behavioral Development*. New York, NY: Academic Press.

Wolfe, C., & Bell, M. (2007). The integration of cognition and emotion during infancy and early childhood: regulatory processes associated with the development of working memory. *Brain and Cognition*, *65*, 3–13.

Yap, S., Ji, L., & Hong, E. (2017). Culture and cognition. In J. T. Wixted & S. Thompson-Schill (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 3), *Language and Thought*. Hoboken, NJ: Wiley.

Zahn-Waxler, C., Park, J., Usher, B., Belouad, F., Cole, P., & Gruber, R. (2008). Young children's representations of conflict and distress: a longitudinal study of boys and girls with disruptive behavior problems. *Development and Psychopathology*, *20*, 99–119.

Zelazo, P., Qu, L., & Müller, U. (2005). Hot and cool aspects of executive function: relations in early development. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young Children's Cognitive Development: Interrelationships among Executive Functioning, Working Memory, Verbal Ability, and Theory of Mind* (pp. 71–93). Mahwah, NJ: Erlbaum.

Zeman, J., & Garber, J. (1996). Display rules for anger, sadness, and pain: it depends on who is watching. *Child Development*, *67*, 957–973.

CHAPTER 10

Moral Reasoning: Theory and Research in Developmental Science

AUDUN DAHL AND MELANIE KILLEN

INTRODUCTION

Moral reasoning plays a central role in psychological functioning throughout the life span. From Aristotle, to Darwin, to Durkheim, and to Freud, theorists have viewed morality as essential to the human condition. Moral philosophers have formulated views of morality that have guided psychological approaches to morality and its development (Appiah, 2005; Gewirth, 1978; Kant, 1785/1959; Nussbaum, 1999; Rawls, 1971; Sen, 2009). Psychological theories of morality address how humans determine the best way to live, form social groups, create norms for regulating social interactions, and challenge social inequalities and unfair treatment of others.

Across the different philosophical formulations of morality, whether the focus lies with virtues, emotions, judgments, or behavior, moral reasoning has been a core component: Reasoning allows humans to go beyond the social abilities of nonhuman primates. Whereas great apes may cooperate with friends and show anger at foes (Tomasello, 2016), only humans apply general evaluative principles to experienced, observed, or hypothetical events from a first-, second-, and third-party perspective (Turiel, 2014).

The role of moral reasoning has been debated in current psychological research on moral judgment. Although the original formulations of moral reasoning in psychological research primarily came from developmental psychology, beginning with Piaget (1932) and followed by Kohlberg (1969), with elaborations from Turiel (1983) and Smetana (2013), the topic has been taken up more recently by experimental psychologists, social psychologists, experimental philosophers, behavioral economists, and evolutionary biologists (Killen & Smetana, 2015). At the center of recent debates is the question of what role, if any, reasoning plays when people make moral judgments.

This chapter discusses evidence about the nature of moral reasoning in children and adults. Although most of the research reviewed comes from developmental psychology, the theories and methods discussed have broader implications for the study of human morality across the disciplines. A key point in this chapter is that moral considerations differ from other evaluative considerations, such as those pertaining to cultural traditions, conventions, group norms, and rituals. These latter forms of regulating group interactions may, at times, include moral considerations, but they are not defined by moral principles. Groups organize themselves in ways that regulate their interactions. The violation of the group norms does not necessarily result in unfair and unequal treatment of others, as we discuss. We argue that developmental changes in children's reasoning within the moral domain, and in their reasoning about conflicts between moral and nonmoral considerations support our proposition that morality is not a subset of group norms. Instead, morality emerges as a set of judgments that are not defined by groups or cultural traditions.

Researchers have used the term "moral reasoning" in diverging ways, in part because morality is studied by so many different fields, within and outside of psychology, and across such a wide age range, from infancy to adulthood. We therefore begin this chapter by defining the moral domain—what moral reasoning is about. We then discuss what kind of evidence is required for inferring that people engage in moral reasoning. We contrast our approach to an intuitionist one, which has argued that much or most of morality does not involve moral reasoning.

In the main part of the chapter, we discuss developmental research on the precursors and early forms of moral reasoning from infancy to preschool age. We then review research on moral reasoning and judgments in group contexts, and specifically reasoning and judgments about challenging topics involving social exclusion, prejudice, bias, and discrimination of others. There are other contexts in which morality is challenged, but we focus on the group—and intergroup—context, given the programmatic research in this area and the timeliness of this area of research for current societal discussions about social inequalities, inequities, and the various forms of prejudice that have arisen in many societies around the globe. Finally, we outline several new directions for research on moral reasoning.

MORAL REASONING

Moral Domain

Children and adults make a large number of evaluative judgments, including personal preferences ("I don't like pizza"), hypothetical imperatives ("If you are going to Los Angeles, you should take Interstate 5"), and interpersonal prohibitions ("Don't hit him!"). Such judgments build on different types of knowledge and experiences and apply to different circumstances, making it necessary to distinguish between types of judgments. Of particular relevance for this chapter is the observation that not all judgments about right and wrong are moral judgments (Smetana, Jambon, & Ball, 2014; Turiel, 1983). For example, declaratives regarding hygiene ("It's wrong to not brush your teeth") or conventions ("One should not wear pajamas to school") are rarely viewed as moral judgments.

In our definition, moral judgments are evaluations based on considerations of others' welfare, rights, fairness, and justice. This definition, rooted in social domain theory (Turiel, 1983, 2002, 2014), is based on philosophical definitions regarding how individuals ought to treat one another (Appiah, 2005; Gewirth, 1978; Nussbaum, 1999; Sen, 2009) as well as an extensive body of empirical research on how individuals evaluate social events and interactions in the world (see section titled "Origins and Development of Moral Reasoning"). The social domain approach has asserted that evaluations of moral considerations (moral domain) are distinct from concerns that pertain to how groups regulate their interactions (societal domain) and those about individual prerogatives (psychological/personal domain). Empirical research conducted in a wide range of societal and cultural contexts (with participants from childhood to adulthood) has supported these assertions (see Turiel & Dahl, in press).

People often perceive events as reflecting moral as well as group or individual concerns (Smetana et al., 2014; Turiel, 2014), however, and this theory is well situated to address what has been referred to as multifaceted domain events. For instance, when faced with decisions about whether to exclude members of an out group, children and adolescents attempt to integrate moral considerations of fairness with societal considerations about how an out-group member might affect group functioning (Killen, Mulvey, & Hitti, 2013). In fact, transitions in children's ability to coordinate conflicting considerations are an important aspect of moral development (Nucci & Turiel, 2009).

However, children and adults do not always give priority to moral concerns when making evaluative judgments in multifaceted situations. That is, by defining morality as concerns with others' welfare, rights, fairness, and justice, we do not claim that these concerns always are, or should be, treated as more important than other evaluative considerations (Tisak & Turiel, 1988; Turiel & Dahl, in press). There are situations in which people give priority to nonmoral concerns, such as conventional concerns with maintaining the smooth functioning on social groups or prudential concerns with avoiding negative personal consequences (Wainryb & Turiel, 1994).

What Is (Moral) Reasoning?

We define reasoning as transitions in thought in accordance with endorsed principles (Adler, 2008; Harman, 1986). An example of a principle of reasoning is: "It is bad to harm others because of the negative intentions to inflict pain on another person." If a person who endorses this principle believes that Tom is harming Henry and therefore judges that "Tom is doing something bad to Henry," then this person has engaged in reasoning. Our definition distinguishes reasoning from other mental processes, such as associations ("This cookie reminds me of one I tasted in my childhood") or spontaneous thoughts ("I have a sudden craving for ice cream"). Although these events are transitions in thought, these transitions do not follow endorsed principles for how one should think or act.

Moral reasoning is, in this view, reasoning based on evaluative judgments pertaining to others' welfare, rights, fairness, or justice. Hence, we distinguish between moral reasoning and other forms of evaluative reasoning, such as reasoning about whether an action violates a social convention or puts the agent in danger. Moral reasoning—the formation of judgments based on moral principles—is not the only form of reasoning relevant to moral judgments. First, reasoning also is involved in the weighing of multiple considerations, such as in intergroup contexts. We return to this type of reasoning later in the chapter (see section titled "Moral Reasoning in Complex Contexts"). Second, reasoning about material facts—termed "informational assumptions"—also can influence moral judgments (Wainryb, 1991). For instance, beliefs about whether corporal punishment teaches children to behave better (an informational assumption) can be changed when faced with counter evidence (children do not learn but acquire antisocial behavior when disciplined with corporal punishment; Turiel, Hildebrandt, & Wainryb, 1991).

Contrary to our proposition that reasoning is central to morality, others have argued that many or most moral judgments are not based on reasoning but on affective, automatic, and unconscious reactions (sometimes called intuitions; e.g. Greene, 2014; Haidt & Bjorklund, 2008). According to this intuitionist view, people regularly form moral judgments based on "gut feelings," such as an aversive affective reaction to the idea of pushing someone or the experience of disgust while watching an action. The differences between our reasoning-based view and the intuitionist views of moral functioning largely consist in two interrelated issues: First, the intuitionist view has adopted a restrictive definition of moral reasoning as relatively slow and effortful. Second, based on this restrictive definition of reasoning, the intuitionist view has claimed that people rarely reason about moral issues.

Proponents of intuitionist views have conceptualized moral reasoning as conscious mental activity; that is, intentional, effortful, and controllable activity (Greene, 2014; Haidt & Bjorklund, 2008). For instance, Haidt and Bjorklund (2008) defined moral reasoning as "conscious mental activity that consists of transforming given information about people in order to reach a moral judgment" (p. 189). They contrast such moral reasoning with "moral intuitions," which are defined as "the sudden appearance in consciousness, or at the fringe of consciousness, of an evaluative feeling (like–dislike, good–bad) about the character or actions of a person, without any conscious awareness of having gone through steps of search, weighing evidence, or inferring a conclusion" (p. 188).

The intuitionist view of moral reasoning as consciously and effortfully going through multiple steps of thought differs starkly from our definition of reasoning as the formation of judgments in accordance with endorsed principles. Our definition of reasoning does not require that people consciously go through multiple steps of thought in order to formulate a reason before making a judgment. On the contrary, over the course of development, many forms of moral reasoning may become so well rehearsed that such reasoning appears to happen automatically and without effort, yet it can be applied flexibly to a variety of situations (Pizzarro & Bloom, 2003; Turiel & Killen, 2010). This view of reasoning as sometimes rehearsed, fast, and effortless is consistent with other theoretical approaches to reasoning (Adler, 2008; Frank & Goodman, 2012; Oaksford & Chater, 2001). The effects of rehearsal on reasoning are evidenced by research on expertise. For instance, physics experts solve even simple physics problems much faster than novices and do seemingly with minimal planning (Larkin, McDermott, Simon, & Simon, 1980). This rehearsed aspect of moral reasoning in adulthood emerges out of a long, protracted development from infancy to late adolescence. What appears to be accessed easily in adulthood (e.g., judgments about the infliction of harm or the denial of resources) takes many years of experience, abstraction, reflection, and action over the course of child and adolescence development, as we discuss below (see section titled "Origins and Development of Moral Reasoning").

In our view, the key criteria for determining whether someone engaged in moral reasoning are whether the person can articulate reasons for their judgments when prompted and whether those reasons are consistent with their judgments. For instance, when explaining why it is wrong to hit someone, preschoolers and older children often say that hitting negatively affects the welfare of the victim (e.g., Dahl & Kim, 2014; Killen & Smetana, 1999; Turiel, 2008). In contrast, when explaining why it is wrong to wear a bathing suit to school, children often refer to rules or authority commands (e.g., there is a rule that you cannot wear a swimsuit to school). These reasons are generally consistent with children's pattern of judgments. For instance, most children would say it would not be okay to hit someone else even if teachers said so it was (since the victim is still hurt), whereas it would be okay to wear a bathing suit to school if there were no teacher commands or rules against it (e.g., if the teacher said it was okay then we could do it, like "pajama day"!). By stating that children (and adults) reason about moral and other social issues, we are not claiming that they always consciously go through steps of reasoning prior to making judgments. Rather, we argue that children and adults readily provide and endorse valid reasons for their judgments in most situations when prompted. (Nonetheless, there are contexts that make the straightforward application of moral reasoning difficult, such as when the attributions of intentions are unclear or group identity changes the interpretation of who fully merits fair and just treatment. In these complex contexts, we expect that moral reasoning is sometimes effortful and may involve consciously going through multiple steps of thought.)

Our view differs from the intuitionist view not only by our definition of moral reasoning but also about how common moral reasoning is in the lives of children and adults. Contrary to our claim about the centrality of moral reasoning, intuitionist approaches have claimed that people often, or even typically, do not reason about moral issues (Haidt, 2008). This claim is, to a large extent, based on research purporting to show that people are either unable to provide justifications for their judgments ("moral dumbfounding") or provide post hoc rationalization that do not explain their judgments (Greene, 2014; Haidt, 2001). If it were the case that people frequently were unable to explain their moral judgments, or provided judgments that were inconsistent with their judgments, this would indeed run counter to our claim that people typically reason about moral issues and that their judgments are based on principles they can articulate and endorse. However, the empirical support for moral dumbfounding and post hoc rationalization is highly limited. And, in contrast, there is extensive empirical support for moral reasoning from childhood to adulthood (see Killen & Smetana, 2015).

Although space prevents an in-depth discussion, we note that claims about moral dumbfounding and post-hoc rationalization are based on just a handful studies. In contrast, a large number of studies have shown that children and adults can justify their judgments (see sections titled "Emergence of Moral Reasoning" and "Moral Reasoning in Complex Contexts"). Furthermore, the few studies purporting to show moral dumbfounding have asked people to judge highly unusual situations (e.g., sex between siblings, sacrificing one life to save others), are unpublished (Haidt, Bjorklund, & Murphy, 2000), have asked people to justify differences between judgments rather than the judgments themselves (Cushman, Young, & Hauser, 2006; Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007), or have included only anecdotal reporting of participants' justifications (Wheatley & Haidt, 2005). These limitations call into question the basis by which researchers have denied the role of meaningful and authentic moral reasoning in how individuals make decisions in their everyday lives. To read more about these issues, the reader is referred to discussions elsewhere (Jacobson, 2012; Royzman, Kim, & Leeman, 2015; Turiel & Dahl, in press).

Our definition of moral reasoning leads to two important clarifications that have implications for research on moral orientations and their development: First, individuals may endorse principles of reasoning not endorsed by researchers; and, second, they may reason about several reasonable options, not just a single option.

People May Endorse Principles Not Endorsed by Researchers

Our definition of reasoning recognizes that individuals hold multiple principles of moral reasoning. These principles sometimes enter into conflict, as in moral dilemmas in which principles of individual rights are pitted against the utilitarian principle of maximizing overall welfare (Foot, 1967; Nussbaum, 1999). Whether a person can be said to hold a given moral principle depends on whether the person endorses that principle upon reflection (Rawls, 1971). If, after thinking about it, a person believes that, other things being equal, it is wrong to harm others, we would say that this person holds the principle that it is wrong to harm others (other things being equal). Hence, our definition of moral reasoning does not require that the principle conform to an a priori criterion of valid reasoning, such as the maxims of rational choice theory or the utilitarian principle of maximizing the sum of welfare across all affected parties (Greene, 2014; Jacobson, 2012).

We avoid a reliance on a small set of "a priori" principles for valid reasoning because the unique validity of such a set would be far from self-evident against the backdrop of the relevant philosophical literature (Elqayam & Evans, 2011). We do not see how psychologists can be the judges of which moral principles are valid and which are "irrational." Rather, we propose that researchers accept as reasonable those principles that reflect philosophical criteria endorsed by their research participants upon reflection. Moreover, empirical psychological research provides the basis by which we can validate that individuals value and cherish these principles.

People May Reason About Several Reasonable Options, Not Just a Single Option

Reasoning does not always lead to a single acceptable solution to a problem (Scanlon, 2014; Searle, 2003). On the contrary, reasoning can leave room for arbitrariness and uncertainty. A child may have very good reasons to lie and very good reasons to tell the truth, making it very difficult to make a reasoned choice between the two courses of action. In the so-called "trolley car dilemmas" (Foot, 1967; Thomson, 1976), in which the protagonist has to choose between letting five persons die and sacrificing another life to save the five persons, people readily articulate both reasons for intervening (e.g., maximizing the number of lives saves) and reasons against intervening (e.g., the general prohibition against actively killing others; Dahl, Uttich, Gingo, & Turiel, 2013). Most people appear to find it very difficult to make a judgment about these and other dilemmas of life and death, not because they fail to reason but, on the contrary, because their reasoning does not yield a unanimous judgment about the situation.

We hypothesize that nonreasoning processes will exert the biggest effects when participants are choosing among several "reasonable" options and lack the time or the information to make principled decisions. For instance, Payne, Jacoby, and Lambert (2005) found that racial bias had the largest effect on decisions when participants are forced to respond quickly and therefore act on less accurate perceptions (more ambiguity). The inverse relation between nonreasoned racial bias and ambiguity of the stimulus further illustrates how indeterminacy of reasoning operates. When people lack compelling reasons for choosing one act or belief over another, nonreasoned processes may play a greater role, so to speak as "tie-breakers" (Kihlstrom, 2013). Further, how individuals assign blame is often a result of the misattribution of intentions. In childhood, for example, children who lack "theory of mind" (i.e., recognizing that others have intentions, desires, and emotions different from the self) are more likely to assign blame in situations than are children who have "theory of mind." This is particularly evident when there is ambiguity regarding the intentions of the transgressor, such as in an "accidental transgressor" context (Killen, Mulvey, Richardson, Jampol, & Woodward, 2011).

Similarly, social psychologists have demonstrated that in straightforward contexts, most individuals show egalitarian orientations; racial bias and stereotyping are more likely to be revealed in situations that are complex or ambiguous, suggesting that when individuals are cognitively overloaded, they resort to stereotypic responses when they have difficulty making decisions (Gaertner & Dovidio, 2005).

Our definitions of morality and reasoning provide a theoretical framework for studying the development of moral reasoning from early childhood to adulthood. In the next section we review research on the origins and development of moral reasoning, followed by a focus on moral reasoning in complex contexts, such as those in which prejudice, bias, and discrimination—forms of social inequalities—are salient.

ORIGINS AND DEVELOPMENT OF MORAL REASONING

Precursors of Moral Reasoning in the First Years

Our definition of moral reasoning requires that children express moral judgments based on principled considerations of rights, others' welfare, or fairness. By this definition, moral reasoning is not present in infancy. We argue that infants do not express judgments of right and wrong as defined by principled considerations, that is, considerations that are generalizable, obligatory, and prescriptive. Yet moral reasoning builds on orientations and skills that emerge and develop during infancy. We therefore differentiate the precursors of moral reasoning, which include empathic responsiveness, social understanding, and signs of guilt and shame, from early moral awareness (for a review, see Killen & Smetana, 2015), which we define as the beginning of obligatory and prescriptive judgments about right and wrong (Gewirth, 1978; Rawls, 1971). In our view, a comprehensive developmental account of moral reasoning will discuss not only developmental changes that take place after children have begun to reason about moral issues but also the building blocks in the first years that make the emergence of moral reasoning possible. In subsequent sections, we discuss research on these early building blocks of moral reasoning.

Empathic Responsiveness to Distress

Empathic responsiveness to the distress of another individual—the concern for the well-being of others—is a key aspect of human morality. If people did not care about others' well-being then they could not have a moral sense as we know it now. However, despite its moral significance, full-fledged empathy is not genetically preprogrammed or present at birth but develops gradually through social interactions over the first years. Also, empathy, even in its fully developed form, is not sufficient for the emergence of moral reasoning; nor is it definitional of the moral domain.

There is some evidence that infants react negatively to others' distress soon after birth. In the standard paradigm, researchers assess infants' reactions to recordings of different crying sounds, both the infants' own cries and the cries of other infants or children. Neonates cry more when hearing the cry of another neonate than when hearing a recording of their own cry (Martin & Clark, 1982; Sagi & Hoffman, 1976). However, neonates also cry more at the sound of another neonate crying than at the sound of an older child crying. The latter finding may reflect important differences between empathic distress in neonates and empathic distress in older children and adults. For instance, some have argued that neonatal crying upon hearing other neonates cry may reflect competition for attention rather than empathy (Campos et al., 2008).

Later in the first year and into the second, infants show increasing levels of interest in or concern with others' distress (Davidov, Zahn-Waxler, Roth-Hanania, & Knafo, 2013; Hay, Nash, & Pedersen, 1981; Roth-Hanania, Davidov, & Zahn-Waxler, 2011; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). As an example, in a study by Roth-Hanania, Davidov, and Zahn-Waxler (2011), 8- to 16-month-old infants witnessed their mother simulating distress (e.g., after pretending to hit her finger with a toy hammer). With age, infants showed growing interest in the mothers' distress, such as by looking back and forth between the hurt finger and the mother's face.

However, even in the second year of life, infants' empathic capabilities remain limited. A phenomenon illustrating these limitations is infants' tendencies to use force against others without any provocation or sign of frustration. In one study, most infants in their second year engaged in such acts of unprovoked force against others (Dahl, 2016a). For instance, an infant could walk up to a parent sitting on the floor and hit the parent in the face without any sign of anger. Since these acts happen without signs of distress, they cannot be attributed easily to limitations in infants' ability to regulate their frustration (Hay, 2005; Thompson & Goodvin, 2007). Although it is currently unclear whether infants fail to understand that these acts cause pain to others or whether they are not sufficiently concerned with preventing others' pain, acts of unprovoked force indicate limitations in infants' concern with the well-being of others. The rate of such behaviors appears to decrease late in the second year, suggesting that children begin to avoid actions that cause distress or negative reactions in others unless they are highly motivated to do so. In fact, the rate of provoked acts of force, as during property conflicts, appears to increase throughout the second year (Dahl, 2016a; Hay, 2005).

Over the course of the second and third years, children increasingly seek to relieve the distress of others, including when they themselves have caused the distress (Hoffman, 2000; Svetlova, Nichols, & Brownell, 2010; Zahn-Waxler et al., 1992). Svetlova, Nichols, and Brownell (2010) introduced 18- and 30-month-olds to a distressed experimenter with a specific need (e.g., an experimenter who was demonstrating behavior that indicated that she was very cold [shivering]). The researcher had previously shown the children how a blanket located in the same room made her warm. The 30-month-olds provided the blanket far more readily than did the 18-month-olds.

Yet there is substantial individual and situational variability in children's propensity to comfort persons in distress. Using parental report and direct observation, studies of children around their second birthday have reported that only in about half of everyday instances of distress do these children respond with attempts to relieve others' distress, as by bringing a toy to cheer the person up (Dunn & Munn, 1986; Eisenberg, Spinrad, & Knafo-Noam, 2015; Zahn-Waxler et al., 1992). Other studies have shown systematic individual differences. For instance, Young, Fox, and Zahn-Waxler (1999) found that infants low in emotional reactivity to novel stimuli at 4 months (e.g., videos of brightly colored mobiles with Winnie-the-Pooh characters) showed less concern for a distressed person at 24 months.

In sum, negative reactions to others' distress are seen from birth. From late in the first year until the end of the second year, infants show substantial increases both in overall interest in others' distress and in attempts to relieve others' suffering, although there is substantial intra- and inter-individual variability.

By definition, empathy requires some grasp of another person's experience or situation (Hoffman, 2000). Changes in social understanding thus likely set the stage for several of the changes in empathic responses seen during infancy and toddlerhood just described. Moreover, the limitations in social understanding at the end of toddlerhood place constraints on children's ability to empathize with and respond to others. For instance, 2-year-olds will have difficulties understanding that a person can be sad after a negative event if the person is not showing outward signs of sadness (Pons, Harris, & Rosnay, 2004). Further advances in social understanding are required before children can perceive and respond to such situations adequately. We now turn to the early development of social understanding and preferences, before discussing the emergence of explicit moral judgments.

Understanding and Evaluating What Others Want: Social Cognitive Development in Infancy and Toddlerhood

Moral reasoning and judgment rely on understanding of others' beliefs, goals, and desires. Fully developed moral evaluations involve attribution of intentions to agents (e.g., the attribution of harmful intent). Moreover, actions often are evaluated on the basis of their intended or foreseeable consequences of actions (e.g., the foreseeable [harmful] consequences of making fun of someone's physical appearance). Thus, moral evaluations of actions build on attributions of psychological states to victims or beneficiaries as well as to agents (Gray, Young, & Waytz, 2012; Killen & Smetana, 2015; Wainryb, 1991). The abilities to understand and respond to others' psychological states undergo major transformations in the first 3 years of life. In this section, we briefly review research on these abilities and their limitations.

Within a few weeks after birth, infants' orientations toward people differ qualitatively from their orientations toward objects (Trevarthen, 1979). Around 4 to 8 weeks of age, infants begin to smile in response to social stimuli, as during vigorous tactile stimulation or the presentation of moving faces (Sroufe, 1996). By 3 months, infants also show greater negativity (such as crying) when mothers are unresponsive than during natural, responsive interactions, suggesting that infants at this age expect and want reciprocal interactions with others (Cohn & Tronick, 1983; Tronick, 1989).

A number of studies have investigated infants' sensitivity to intentional actions from early in the first year, typically relying on looking patterns to indicate whether infants' perceive actions as intentional. For instance, Woodward (1998) habituated 5- and 9-month-olds to seeing a hand reaching for one of two toys. In the test trials, the location of the toys was flipped. Infants saw the hand either reaching for the original toy in the new location (new movement but same goal as before) or reaching for the new toy in the original location (same movement but new goal). Infants looked longer when the hand reached for the new toy in the original location than when it reached for the original toy in the new location, suggesting that infants expected the hand to seek a particular object rather than repeating the same physical movement. (Findings were only marginally significant for 5-month-olds.) When a rod rather than a hand was "reaching" for the toy, infants looked longer at the new movement/old goal display, suggesting that they did not view the rod as intending to reach the ball. Other studies have found that infants may perceive actions as intentional already at 3 months if given experience with holding objects (Sommerville, Woodward, & Needham, 2005; Woodward, 2009). In fact, infants' understanding that others have goals, beliefs, and desires at age 7 months (goal encoding) and at 18 months (implicit false belief understanding) predicts children's morally relevant theory of mind (an understanding of an accidental transgressor's moral intentions) at age 5 years (Sodian et al., 2016).

There is also evidence that infants, from early in the first year, visually prefer "prosocial" puppets (acting in ways that are conducive to others' goals) over "antisocial" puppets (acting in ways that keep others from reaching their goals). Hamlin and Wynn (2011) showed 3- and 5-month-olds interactions among three puppets. A neutral puppet rolled a ball to the two other puppets. The "antisocial" puppet kept the ball, whereas the "prosocial" puppet returned it to the neutral puppet. Three-month-olds preferred (i.e., looked more toward) the prosocial than the antisocial puppet, and 5-month-olds were more likely to reach toward the prosocial than the antisocial puppet (Hamlin, 2014; Hamlin, Wynn, & Bloom, 2007).

Infants' preferences for looking or reaching toward certain agents likely influence their subsequent development regarding agent–action relationships. These preferences, however, differ from moral judgments observed later in childhood. First, it is unclear whether these preferences reflect moral evaluations. Instead, it may be that these preferences simply reveal desires to interact with prosocial agents. Further indicating that infants' preferences may not reflect moral evaluations, several studies have shown that infants' preferences also are guided by events that are morally irrelevant, such as whether the agent shares food preferences with the child or previously has imitated another agent (Hamlin, Mahajan, Liberman, & Wynn, 2013; Powell & Spelke, 2014). Moreover, young children's decisions to approach one agent rather than another appear to be based on preferences as opposed to categorical evaluations. For instance, in some studies, toddlers and preschoolers observed one agent acting helpful toward a neutral agent and another agent acting in an uncooperative or destructive way (Dahl, Schuck, & Campos, 2013; Vaish, Carpenter, & Tomasello, 2010). Although children preferred to help the helpful agent in a subsequent task, most were willing to help the uncooperative or destructive agent if the helpful agent was unavailable. Thus, infants' third-party reactions to violations (e.g., hindering actions) may be relative (a preference for one over the other) rather than categorical (a categorically negative evaluation of one agent, as in a moral condemnation; Dahl, 2014).

Late in the first year and into the second, a strong sensitivity to others' intentions and desires is evident in a variety of infants' actions. Infants are increasingly skilled at following the point and gaze of another person, even when the person is attending to objects outside the infants' field of view (Butterworth & Jarrett, 1991). By the middle of the second year, infants can not only follow but also direct others' visual attention toward objects (Carpenter, Nagell, & Tomasello, 1998).

From around the first birthday, infants are able to use their understanding of others' intentions to join and contribute to others' activities. Warneken and Tomasello (2007) presented 14-month-olds with an adult researcher who needed help (e.g., because she dropped a marker on the floor and could not reach it). The majority of infants helped at least once (e.g., by handing back the dropped the marker). Later in the second year, children help in more complex tasks and sometimes even without explicit signals of need from the experimenter (Warneken, 2013; Warneken & Tomasello, 2006): Warneken (2013) found that 24-month-olds would help an experimenter who unknowingly knocked an object off the table even before the experimenter noticed that the object had fallen down.

Infant helping and participation also take place in family homes, not just in the laboratory (Dahl, 2015; Dunn & Munn, 1986). In fact, several pieces of evidence suggest that infant helping and participation may be facilitated by adult encouragement during every day social interactions. Dahl (2015) used maternal reports and direct observations of everyday interactions to document the social context of helping in the everyday life of U.S. middle-class families. Most infants engaged in at least some forms of helping from around the first birthday. Moreover, longitudinal data showed that helping rates were positively associated with caregiver encouragement of helping on previous observation points (Pettygrove, Hammond, Karahuta, Waugh, & Brownell, 2013; Waugh, Brownell, & Pollock, 2015).

An experimental follow-up further suggested that adult encouragement plays a crucial role in the early development of infant helping (Dahl, Satlof-Bedrick, et al., 2017). In this study, infants aged 13 to 18 months witnessed an experimenter accidentally drop objects, such as a pen, onto the floor and unsuccessfully reach for them. A second experimenter was also present and played with the infants between each trial. Half of infants were assigned to receive encouragement (e.g. "Do you want to help her?," "Do you want to hand her the pen?") and praise ("Great job! Thank you for helping!"), while the other half served as a control group and received no encouragement or praise from the experimenters. Among the younger infants in the study, encouragement and praise doubled helping rates, both on trials when infants were encouraged and praised and on subsequent trials without encouragement or praise. Coding of infants' looking behavior indicated that the experimental manipulation did not merely help infants notice the experimenter in need, as all infants looked toward the reaching experimenter. In contrast, the older infants in the study appeared to have mastered this simple helping task without the need for adult support. Moreover, older infants' helping rates were higher than those of younger infants and were unaffected by the experimental manipulation (Warneken & Tomasello, 2013). However, it is possible, and even likely, that adult encouragement and other forms of support may facilitate children's helping on more complex tasks (Hammond & Carpendale, 2015; Rogoff, 2003).

In the second year, infants also guide their behaviors by emotional and other signals about what others do not want. In a classic study on social referencing, Sorce, Emde, Campos, and Klinnert (1985) found that 12-month-olds were more likely to avoid an apparent 30 cm "cliff" (covered by transparent glass) when their mothers displayed a fearful or angry facial expression than when their mothers displayed joy or interest. Gradually, infants also come to realize their desires are not always shared by others. In a study by Repacholi and Gopnik (1997), 14- and 18-month-olds first observed whether an adult expressed disgust or joy when eating broccoli or Goldfish crackers. The infants then had the option of giving either broccoli or Goldfish crackers to the adult. The younger children tended to give whichever food type they preferred for themselves, while the older children tended to give the type of food over which the experimenter previously had expressed joy.

Negative signals from others are especially common after prohibited or unwanted behaviors in the everyday lives of infants and their families. Naturalistic studies have found that conflicts about prohibited behaviors can occur 10 or more times per hour in the second year (Dahl, 2016b; Kuczynski, Kochanska, Radke-Yarrow, & Girnius-Brown, 1987; Power & Parke, 1986). These conflicts provide young children with information about not only other people's expectations but also the nature of those expectations. From early in the second year, if not before, there are systematic differences in how caregivers respond to different types of infant violations. Several studies have found that when infants hit, bit, kicked, or otherwise used force against others (a moral violation), caregivers responded with heightened anger, more physical interventions, and more references to the consequences of such acts for other people (Dahl, 2015; Dahl & Campos, 2013; Dahl, Sherlock, Campos, & Theunissen, 2014; Smetana, 1989). In contrast, when infants created inconvenience, such as by throwing food on the floor, or when they do something that could affect their own welfare, such as climbing on a couch, caregivers responded with more distraction (e.g., drawing infant attention to a toy), more positive tones of voice (e.g., to comfort infant after prohibition), and more compromising (e.g., letting infant engage in the prohibited behavior for a limited period and under adult supervision).

Children become increasingly aware of others' negative reactions to their behaviors over the course of the second year. A clear sign of such awareness is children's tendency to look at a parent and smile, or even elicit the parent's attention, before engaging in a prohibited behavior (e.g., before approaching a prohibited kitchen cabinet). The first signs of such anticipation are reported around the first birthday, but the behavior appears to become increasingly common during the second year (Bretherton & Bates, 1979; Dahl & Freda, 2017; Dunn & Munn, 1985). Relatedly, longitudinal naturalistic research by Dunn (1988) described how children became adept at teasing over the course of the second year by acting in ways they knew would upset their older sibling.

When receiving prohibitions, children become increasingly likely to acknowledge, rather than simply ignore, such prohibitions during infancy and toddlerhood. However, children's increasingly common acknowledgments of parental commands may take the form of compliance, negotiation, or refusal (Dahl, 2016b; Kaler & Kopp, 1990; Klimes-Dougan & Kopp, 1999; Kuczynski et al., 1987). In short, improvements in children's social understanding during the second year go along with an increasing awareness of what others expect of them, but this awareness does not by itself lead children to accept and meet those expectations. An additional, crucial component is children's evaluation of their actions and the consequences of those actions.

Children's negative evaluations of their own transgressions may be based on a concern with not upsetting others or with a genuine adoption of a particular rule, such as the prohibition against harming others. Children's concerns with not upsetting others are indicated, in part, by the precursors of guilt and shame. Importantly, the precursors of guilt and shame are not by themselves signs of moral reasoning but are, at most, early steps toward genuine moral judgments and reasoning. We briefly review the early development of guilt and shame in the next subsection. Children's endorsement of rules is seen most unambiguously in their third-party evaluations of interactions between others, for instance when one person is hitting or stealing from another. Evidence of such categorical third-party judgments, in the form of protests or verbal judgments, has been found in the third year (as we discuss in the section titled "Emergence of Moral Reasoning: Early Preschool Years").

Early Roots of Guilt and Shame

The earliest signs of self-evaluation have been reported in the second half the second year (Barrett, 2005; Cole, Barrett, & Zahn-Waxler, 1992; Stipek, Recchia, McClintic, & Lewis, 1992). Research on the development of guilt and shame—two key affective correlates of self-evaluation—often has used the so-called "broken toy" paradigm (Barrett, 2005; Cole et al., 1992; Kochanska, Gross, Lin, & Nichols, 2002). In this paradigm, children get to play with a toy said to be important for the adult researcher. The toy is rigged so that when the children touch it in the researcher's absence, the toy's leg breaks. Barrett (2005) found that, at 17 months, a large proportion of children engaged in guilt-related behaviors (e.g., drawing the parent's or experimenter's attention to the broken leg) or shame-related behaviors (e.g., avoiding the experimenter when she returns). Attempts to repair and tell the experimenter, as well as gaze aversion, are even more common among 2-year-olds than among 17-month-olds in the broken toy paradigm (Barrett, 2005; Barrett, Zahn-Waxler, & Cole, 1993).

Importantly, the reported attempts at reparation, communication, and avoidance in the broken toy paradigm may or may not reflect children's negative self-evaluations of their own actions. These reactions could equally well indicate a concern with the researcher's anticipated negative reaction or curiosity about the broken object (Kagan, 1981).

There is limited evidence for when children begin to apply unambiguous negative evaluations of their own moral and nonmoral violations (Eisenberg, 2000). Anecdotal evidence from a study by Zahn-Waxler and Kochanska (1990) suggest that at least some children apply negative evaluations to themselves already in the second year and even apologize for hurting others. Still, signals of guilt based on a negative self-evaluation, as described in these anecdotes, appeared clearer and more robust in the third year. Stipek, Gralinski, and Kopp (1990) found that, although some mothers reported that their infants expressed self-evaluation and negative reactions at their own transgressions in the second year, the proportion of reporting such behaviors in their infants increased dramatically into the third year. These findings are consistent with theoretical proposals by Mascolo and Fischer (2007), who argued that children in their third year become capable of expressing guilt or remorse even in the absence of explicit negative reactions from others.

The development of guilt and shame continues into the preschool years and beyond (Mascolo & Fischer, 2007). Recent studies found children to be more motivated to help another person after causing accidental damage to the person's property than when they had not caused such damage (Hepach, Vaish, & Tomasello, 2017; Vaish, Carpenter, & Tomasello, 2016). Another study found that 2-year-olds who showed guilt-like behaviors in a broken toy paradigm were more helpful toward the "victim" in a subsequent situation than children who showed shame-like (avoidant) behaviors (Drummond, Hammond, Satlof-Bedrick, Waugh, & Brownell, 2017). Whatever the nature of children's self-evaluation, these studies suggest that young children's perceptions of their own responsibility for events influence their actions toward others.

In sum, the first years are a period when infants make dramatic advances in their empathic responsiveness to distress and in their social understanding. At the same time, there is no clear evidence that infants make moral evaluations of right and wrong based on principles about well-being, rights, and justice. The stage is set for a major moral transition taking place in the early preschool years.

Emergence of Moral Reasoning: Early Preschool Years

The third year of life reveals the emergence of moral reasoning with the onset of verbal explanations that include prescriptive, obligatory judgments regarding specific acts of harm, the denial of resources, and victimization. At this age, children begin to express judgments of right and wrong and protest against rule violations (Rakoczy, Warneken, & Tomasello, 2008; Rizzo & Killen, 2016; Smetana & Braeges, 1990; Smetana et al., 2012). Children begin to articulate justifications for their judgments in the third year of life. Their reasons are not consistently applied to all judgments but are reliably codified and indicate principled moral considerations about welfare, rights, and fairness. In one paradigm, children are asked about a series of hypothetical situations involving moral and conventional rule violations (e.g., a child hitting or stealing from another child). By the end of the third year, most children judge it as wrong to harm or steal from others even if teachers or parents were to say it was permissible and even if it happened in another school where there was no rule against it (Smetana et al., 2012). In contrast, children at this age tend to judge violations of social conventions as permissible if adults said these acts were permissible or if there was no rule against them, revealing an underlying distinction between moral and social-conventional transgressions.

These judgments indicate that children view moral rules, but not social conventions, as based on intrinsic features of the actions (e.g., harm to a victim) rather than proscriptions from adults. In a different paradigm, children witness puppets committing transgressions against each other. For instance, 3-year-olds protested more when one puppet destroyed the drawing made by another puppet than when the puppet destroyed his own drawing (Vaish, Missana, & Tomasello, 2011; see also Schmidt, Rakoczy, & Tomasello, 2013).

Between the third and the fourth birthday, children's distinctions between moral and other events becomes increasingly robust (Smetana, 2013). At this age, children distinguish hypothetical moral violations not only from conventional violations but also from prudential violations (pertaining to the agent's own welfare) and personal issues (issues under personal jurisdiction, e.g., choice of clothing; Dahl & Kim, 2014; Killen & Smetana, 1999; Nucci & Weber, 1995; Tisak, 1993). Their principled moral concerns with others' welfare, rights, and fairness are reflected both in their patterns of judgments (such as whether the action would be permissible in the absence of a rule) and also in their justifications. When asked to justify why it is wrong to harm or steal from others, 3-year-olds, as well as older children, refer to consequences to the victim ("It hurts him!") or property rights ("It's his truck!"; Dahl & Kim, 2014; Killen & Smetana, 1999; Nucci & Weber, 1995).

These judgments reveal children's spontaneous forms of reasoning in response to semistructured probes in which counterintuitive premises are described to children for their evaluation (e.g., "What if the teacher said it was all right to hit someone? Then would it be all right?"). The value of the counterintuitive premise is that it reveals whether children are using an underlying principle to explain their judgment (e.g., in response to the teacher premise: "It would still be wrong because someone would be hurt and how will they feel that you did that?"). Children will reject authority mandates and punishment as the reasons for why it is wrong to inflict harm on others and instead refer to the pain experienced by a victim.

Preschoolers' judgments and reasoning about moral issues appear to build on the empathic tendencies and emotional skills that develop during the early years. For instance, studies have found positive associations between preschoolers' empathic responsiveness to others' distress and how negatively they view moral transgressions (Ball, Smetana, & Sturge-Apple, 2017) and their reasoning about and engagement in helpful actions (Miller, Eisenberg, Fabes, & Shell, 1996). Other research has shown that preschoolers reason about, and sometimes regret, their own harmful actions toward others (Wainryb, Brehl, & Matwin, 2005). Further, children starting at 3 years of age recognize the unfairness as well as the harm to others' welfare when someone is denied necessary resources (Rizzo & Killen, 2016).

The conceptual distinctions between moral and other violations are reflected in children's social interactions with one another as well as their responses to hypothetical scenarios (Turiel, 2008). In everyday social interactions, preschoolers, as well as adults, respond differently to moral violations than to conventional violations (e.g., referencing consequences to the victim in response to moral violations but not in response to conventional violations; Killen & Smetana, 1999; Nucci & Turiel, 1978; Smetana, 1989). In sum, preschoolers show a principled (general) concern with the protection and the promotion of others' rights and well-being that is reflected in their judgments as well as their actions.

MORAL REASONING IN COMPLEX CONTEXTS

Children's moral reasoning in the third and fourth years of life are revealed in fairly straightforward contexts (transgressions such as the infliction of harm and the denial of resources, which comprise the most common violations to moral principles). The social world of older children becomes increasingly complex, and the major developmental changes are seen less in dealing with straightforward issues (where we see development in the early years) but in dealing with multifaceted issues. For example, social relationships increasingly are comprised of nonparental adults (at school, in the neighborhood, and the community), peer groups (beyond dyads and triads), complex friendships (best friends, friends, acquaintances, antagonistic peers), and strangers. (For a review, see Rubin, Bukowski, & Parker, 2006.) Along with an expanded social world, children's social cognitive abilities change dramatically. Children's social cognitive capacities include mental state knowledge ("theory of mind"; Wellman & Liu, 2004), judgments about intentional states (Turiel, 2002), and intergroup attitudes and relationships (Rutland & Killen, 2015).

The complexity of the social world and the increase in social cognitive judgments make the task of the application of moral reasoning to everyday life multifaceted and nuanced. How does children's understanding of the group that they belong to guide their decisions about whom to include and whom to exclude? To what extent do children take social relationships into account when making decisions about how to divide resources? Over the past several decades, these questions have been addressed by researchers pursuing how moral reasoning changes over the course of child and adolescent development (Killen, Elenbaas, & Rutland, 2015; Mulvey, 2016).

In many situations, decisions involving moral considerations create dilemmas. Research on moral judgment has investigated how children and adolescents weigh multiple considerations when evaluating such dilemmas. To do so requires identifying other nonmoral compelling considerations. As mentioned, one aspect of children's worlds that changes after early childhood has to do with the onset of intergroup attitudes and relationships (Killen & Rutland, 2011; Nesdale & Flesser, 2001). This context creates some of the most challenging situations for the application of morality because out-group attitudes can transform into discrimination, social exclusion, and bias, which reflect some of the most atrocious forms of moral transgressions in adulthood (R. J. Brown & Gaertner, 2001). Thus, attention to the origins of these contexts for moral reasoning in childhood is warranted. In the next subsection, we provide a review of how children weigh multiple decisions in morally relevant intergroup contexts.

Morality and Group Identity

One of the compelling considerations that make moral decisions complex has to do with the role of groups and group identity in social life (Rutland, Killen, & Abrams, 2010). Early on, children identify with and affiliate with groups, forming attachments that provide a buffer to the complexity and often discomfort associated with a multitude of social expectations (Nesdale & Flesser, 2001; Verkuyten & Thijs, 2006). Groups often are organized by highly perceptually salient features, such as gender, race, and ethnicity. In addition, groups are formed by shared interests and activities. Not unlike the adult world, however, children form in-group preferences as they affiliate with groups (Bennett & Sani, 2004). In morally relevant contexts, in-group preferences can manifest as inclusion preferences (preferring in-group peers in situations involving opportunities) as well as resource preferences (allocating more resources to in-group peers than to others). In-group preferences that turn into out-group dislike form the basis of prejudice as well as discrimination and bias (Nesdale, Durkin, Maass, & Griffiths, 2005).

Thus, these contexts are different from straightforward moral transgressions where the challenge to acting in a way consistent with one's moral reasoning is the opportunity for selfish gain. Here the motivation is to preserve the in-group at a cost to treatment of the out-group. In this case, there is potentially a nonselfish gain for preferring the in-group, one that provides a justification that can be interpreted as legitimate, such as diverting resources to a member of an in-group instead of an out-group (but not to the self). Nonetheless, the implication of such acts has been viewed as wrong from a moral viewpoint because it violates expectations of impartiality and fairness. How children make moral decisions in the context of intergroup relationships has been the focus of much recent research in developmental science, demonstrating the age-related changes that exist for these types of decisions.

Different theories have been proposed, identifying how cognitive, emotional, motivational, and relational changes explain, in part, the age-related changes in identifying with groups and making moral judgments. Studies examine implicit attitudes and biases—responses that people may be unaware of and that have negative consequences to others. At the same time, there has also been a robust body of research on the explicit social and moral reasoning that children and adolescents provide to explain their evaluation of intergroup contexts, such as interracial ones or situations in which others are viewed as out-groups due to their gender, sexual orientation, different cultural membership, or immigrant status. These explicit judgments reveal areas of inconsistencies in moral reasoning created by the salience of group identity and the pervasiveness of messages in cultures regarding maintaining status hierarchies, power, and the status quo (Ridgeway, 2013). Thus, in the next subsection, we report on the developmental trajectories in moral reasoning in group contexts and what the data reveal for how children weigh these complex considerations. We organize the report of the empirical studies by starting with morality and intergroup social exclusion, followed by morality and intergroup allocation of resources and social inequalities.

Morality and Social Exclusion

Social exclusion from groups is an event that occurs frequently in social life. In childhood, these exclusion events often are extremely salient with long-term negative consequences (Killen & Rutland, 2011). With age, children understand that there are many contexts in which social exclusion is justified to make the group work well. For example, for a swim team to be competitive, the team has to exclude those swimmers who are too slow to help the team win. Children learn early that social exclusion is often necessary, even when the excluded individual might feel disappointed. There are also contexts, though, in which morality enters into the decision (not just the outcome), such as when peers are excluded for reasons based on group membership, such as race, ethnicity, gender, and nationality. In these cases, the decision is unfair and violates basic norms about equal treatment. Children are very much attuned to issues of equality, but in the context of social groups, these decisions become difficult. In the case of intergroup exclusion (excluding a swimmer because of ethnicity, not talent), morality is pitted against group membership. With age, children become able to apply their moral knowledge to these contexts.

One line of research investigated children's evaluation of contexts in which one group of children excludes a peer from joining the group when stereotypes are activated (Rutland & Killen, 2015). As one example, when young children were asked whether it was okay to exclude someone who did not fit the stereotypic expectations of an activity-based peer group (playing dolls or trucks), the majority of children ages 3 to 5 years (87%) indicated that it would be unfair ("Dolls are for everyone—that's not fair"; "Girls like trucks too, and they will feel sad if the boys don't let them play"; Killen, Pisacane, Lee-Kim, & Ardila-Rey, 2001). Despite the fact that children are consistently exposed to gender stereotypes (Ruble, Martin, & Berenbaum, 2006) and that children have high knowledge for what types of activities are associated with gender (Liben & Bigler, 2002), children view social exclusion based on gender-stereotypic expectations as unfair and wrong.

Yet research has shown that stereotypes often are activated when situations are complex or ambiguous. Inclusion decisions, for example, are often more complex than exclusion decisions when the inclusion choice is to choose between two individuals. In the study just described (Killen et al., 2001), children also were asked whom to include in their doll-playing or truck-playing peer group. (Should the girls' group pick the boy *or* the girl for doll-playing? Should the boys' group pick the girl *or* the boy for truck-playing?) In this condition, children cited group identity (girls want the girl to join) or stereotypes (boys do not like dolls) to justify their inclusion decision of the child who fit the stereotype (Killen et al., 2001), even when they took a fully moral position in the straightforward exclusion condition. Children recognized the unfairness pertaining to gender exclusion from activities but also were influenced by stereotypic expectations when making decisions that were complex (such as whom to include).

To investigate how group identity plays a role on children's moral judgments after the preschool years, research has examined group dynamics in the context of moral decision making. The term "group dynamics" refers to group norms and the role of group loyalty (Abrams, Rutland, Pelletier, & Ferrell, 2009). Using the group dynamics framework (Abrams & Rutland, 2008), studies have examined whether group identity is defined by children as group membership (loyalty to one's group makeup defined by gender, race, nationality) or group norms (traditions and moral values held by the group).

Previous research has shown that, with age, children give priority to group norms (traditions and moral values) over group membership (gender, nationality). When asked about whether one's own group (boys or girls) would be likely to exclude an in-group member who deviated from (or rejects) the norms of the in-group, with age, children expected that the group would not want to exclude that in-group member; group loyalty matters (Killen et al., 2013). For example, children from 9 to 13 years of age were assigned to actual groups (asked to create a group name and logo) and informed that their group had an unequal norm (preferring to divide resources to advantage their own group) and an out-group had an equal norm (preferring to divide up resources equally between the in-group and the out-group). Children and adolescents gave priority to adhering to the norm of equal allocation (even if it meant rejecting a member of their own group and including the out-group member); equality matters even at the cost of allegiance to the group norm (Killen, Rutland, Abrams, Mulvey, & Hitti, 2013).

This set of judgments involved a complex decision-making process because children had to reject a member of their own group who advocated against equality; equality trumped group loyalty. Although the younger children focused on the moral norm of equality, they also were willing to exclude an in-group member who did not support the norm; older children were not as willing to exclude this member, citing group identity as their reason (Hitti, Mulvey, Rutland, Abrams, & Killen, 2013).

Moreover, children do not treat different types of group identity the same in morally relevant contexts. When comparing gender identity, for example, with school identity (identity based on school affiliation—your school versus a rival school), children were less concerned with group identity in the gender context. For the school-based identity, children were likely to support their in-group member who wanted to give more resources to their own group than to the rival school group (Mulvey, Hitti, Rutland, Abrams, & Killen, 2014). These findings reveal how different aspects of group identity and group norms are taken into account when making decisions about inclusion and exclusion as well as resource allocation.

To further probe the tension between morality and group norms, children's evaluations of group dynamics in a gender context in which stereotypes were highly salient, such as football and ballet, were investigated (Mulvey & Killen, 2015). The goal was to determine whether participants, ages 9 to 14 years, thought that their peers would support an in-group member who challenged a gender-stereotypic activity (such as supporting a group member who asked the girls' group to try football instead of ballet). With age, participants were less likely to expect that their group would support the challenger, even though they personally supported this type of resistance, and they were more likely to expect that anyone who challenged the group would be excluded from the group. Children also evaluated this lack of support as unfair but fitting in with group expectations.

Thus, with age, children viewed the lack of support for challenging gender-stereotypic expectations as unfair but also as part of how groups maintain their identity. The strength of group identity in adolescence is well documented (Horn, 2012; Thijs, Verkuyten, & Grundel, 2014), and how it intersects with moral judgments provides new information regarding the contexts in which group identity maintains its salience. This knowledge is important as it provides information on how to enable groups to be more inclusive. The term "social exclusion" refers not only to the exclusion of peers from social groups; it also refers to the exclusion of individuals from opportunities that are necessary for healthy well-being. In these situations, the denial of access to resources and opportunities provides another complex moral context in which children's moral reasoning emerges early in development.

Moral Reasoning and the Allocation of Resources and Social Inequalities

Research on children's allocation of resources has investigated the claims for resources that children view to be important when determining what constitutes a fair and equal allocation (Kanngiesser & Warneken, 2012; Kenward & Dahl, 2011; Shaw & Olson, 2012). With a few exceptions, research has measured children's preferences and choices (for how to distribute resources) and only a few studies have examined children's moral reasoning (Blake & McAuliffe, 2011; Damon, 1977; McGillicuddy-De Lisi, Daly, & Neal, 2006; Schmidt, Svetlova, Johe, & Tomasello, 2016). Studies that have examined moral reasoning have demonstrated that, for the most part, young children focus on strict equality; with age, children use reasons based on merit and effort. How children coordinate their concepts of equality and merit in different allocation contexts is not well understood.

To test whether younger children take both equality and merit into account, children 3 to 8 years of age were asked to distribute resources that were necessary (need to have to stay healthy) or "luxuries" (fun to have and play with but not necessary; Rizzo, Elenbaas, Cooley, & Killen, 2016). This distinction was investigated to examine whether children would use moral reasoning based on a concern for others' welfare in the necessary condition, along with merit when the recipient was deserving of the resources. Most of the prior research with young children has focused on luxury resources (the distribution of cookies, candy, and stickers) and found that children first use equality reasoning and then use moral reasoning based on merit. In the Rizzo et al. (2016) study, children distributed necessary resources (e.g., such as medicine) equally to a hardworking or lazy character and used reasoning based on others' welfare. Children distributed luxury resources differently, however, giving more to a hardworking character than to a lazy character, using reasons based on effort and merit. Thus, children as young as 3 to 8 years of age evidenced three types of moral reasoning in this context: equality, equity, and others' welfare. A novel dimension of the findings was the use of moral reasoning about others' welfare by young children when considering resource allocation, given the predominant focus on whether children use reasoning based on equity. In fact, children give priority to considerations for others' welfare over equity when resources are necessary for healthy development.

Turning the focus more specifically to the large group context of necessary resources, Elenbaas, Cooley, Rizzo, and Killen (2016) investigated how children distribute necessary resources to groups of children who were disadvantaged (i.e., those who were lacking access to necessary resources, such as school supplies). Children gave more resources to the disadvantaged children than to those who were not disadvantaged using reasoning based on a moral concern for equity and others' welfare. Rectifying inequalities involves complex moral reasoning: To ensure fairness, one has to distribute resources "unequally" to "level the playing field." Although the majority of all children gave more resources to the disadvantaged children, there were also age-related differences in whether in-group bias was displayed. Younger children gave more school supplies when their own group (by race) was disadvantaged than when the out-group was disadvantaged. Older children did not display a bias and, in fact, gave more resources to the societally consistent disadvantaged group (African American) than to the other group, using reasoning that referred to past inequality.

Thus, with age, children rectified the inequality and explained their decisions using moral reasoning, such as references to others' welfare, and social equality. In the study by Elenbaas and colleagues (2016), social knowledge about the factors that contribute to social inequalities were related to children's distribution behavior and their moral reasoning. Thus, children display moral reasoning in complex situations involving social exclusion and resource allocation, even when their moral judgments are challenged by competing claims of group identity, group norms, and societal messages about maintaining the status quo.

In sum, moral reasoning emerges during the preschool period and is applied to a wide range of social contexts, including those that concern social exclusion and resource allocation. Children and adolescents apply their moral reasoning to many complex situations, including those involving parent–adolescent conflict (Smetana, 2011), civil liberties (Helwig, Ruck, & Peterson-Badali, 2014), cultural conflict (Wainryb & Pasupathi, 2009), and sexual identity (Horn, 2012).

CONCLUSION

Moral reasoning is evident in early childhood. The roots are established with early social orientations, empathetic understanding, and moral awareness. Children apply moral reasoning to basic prototypic moral transgressions and understand that not all rule transgressions are the same. Yet as social life becomes highly differentiated with multiple arenas of social relationships (school, family, neighborhood, community) and with the development of different areas of social knowledge (intentionality, group identity, group norms), the application of moral reasoning to social events and interactions is often complicated. Nonetheless, children's moral reasoning continues to be robust, stable, and consistent as they encounter multifaceted situations. Throughout childhood, children not only use moral reasoning in complex situations but challenge social inequalities and inequities. When children advocate for conformity to group norms at the expense of fairness, they often justify these decisions based on conventions, customs, and traditions of the group (rather than resort to selfish orientations). Yet even as groups have a powerful influence on moral decisions, children maintain their concerns for the fair and just treatment of others.

There are also many situations in which children do not challenge moral transgressions that they witness or hear about. Explanations for a lack of acting on one's moral viewpoints include fears of retaliation and exclusion, the misattribution of intentionality of others, the salience of group loyalty, and an uncertainty that intervention will be effective. The role of adults in enabling children to understand connections between acts and consequences, to disentangle moral and nonmoral elements in situations that are multifaceted, and to support children's desires to challenge inequality and unfairness is essential.

This chapter began with definitions of morality (issues of others' welfare, rights, fairness, and justice) and moral reasoning (the formation of judgments on the basis of endorsed moral principles). By these definitions, not all evaluative issues are moral, and not all processes leading to moral judgments count as reasoning. Indeed, several parts of this chapter were dedicated to describing the integration of moral and nonmoral (e.g., conventional) reasoning and interactions between reasoning and nonreasoning processes. We now suggest some key areas of future research on the development of moral reasoning and its relation to other aspects of psychological functioning.

First, how do children adopt new moral principles? Not all principles emerge at the same time. For instance, although property rights are seen in preschoolers, there is less evidence that preschoolers believe that people have other types of rights such as those pertaining to autonomy and free speech, as examples (Helwig et al., 2014; Rossano, Rakoczy, & Tomasello, 2011; Schmidt, Rakoczy, & Tomasello, 2013; Tisak, 1993; Vaish, Missana, & Tomasello, 2011). What kinds of thoughts and experiences lead children to endorse new rights? Answering these developmental questions will require not only interviews and laboratory experiments but also research into children's everyday experiences (e.g., through naturalistic observations; Dahl, 2017; Turiel, 1983; Willems, 1967).

Second, what leads children and adults to change their ways of coordinating different evaluative principles? For instance, how do children come to concern themselves increasingly with historical and societal inequalities in deciding how resources should be allocated? (Elenbaas et al., 2016). The world of childhood is both vertical and hierarchical as children develop concepts of equality, on one hand, and create social status categories that are associated with power and entitlement, on the other hand (C. S. Brown & Bigler, 2005). Children make inferences about what they witness in their peer culture, leading them to accept and reject power hierarchies and status, but doing so is not easy (Mulvey et al., 2013; Nesdale & Flesser, 2001). Challenging the group has a high cost, including potential exclusion from the group. With age, children transfer their knowledge about the peer world to the larger societal world with its traditions and norms that reinforce hierarchical and, at times, unfair treatment of others. This knowledge enables children to make decisions that will either rectify or perpetuate social inequalities that stem from the broader culture, such as those based on gender, race, ethnicity, and nationality (Elenbaas & Killen, 2016).

Third, how do reasoning and nonreasoning processes jointly influence judgments? Much has been made of findings suggesting that manipulations of incidental features (e.g., whether research participants are sitting in a dirty or smelly room) influence evaluative judgments (Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005). However, the effects of incidental disgust on evaluative judgments are generally small and rarely compared the effects of endorsed principles (Kayyal, Pochedly, McCarthy, & Russell, 2015; Kihlstrom, 2013; Landy & Goodwin, 2015). The effects of incidental disgust typically involve making negative evaluations slightly more negative rather than, for instance, making positive or neutral judgments negative (Pizarro, Inbar, & Helion, 2011). Moreover, other "gut" reactions, such as fear, do not necessarily restrict moral action or moral reasoning, as shown by acts of disobedience designed to create social justice (Appiah, 2005; Nussbaum, 1999). In contrast, the presence of an action that violates a moral principle typically is sufficient, and even necessary, in order for children and adults to judge an action as wrong (Killen & Smetana, 2015).

Fourth, what types of social experiences are motivating for children for acting on their moral reasoning in a range of situations that they confront in social life? Developmental science often identifies essential experiences in broad categories, such as peer or adult–child relationships. More recently research has demonstrated that specific types of peer relationships are important to enhance moral judgments, such as cross-group friendships (Tropp & Prenovost, 2008). But what specific aspects of these relationships compel children to challenge wrong deeds by others? Research has focused on whether children attend ethnically homogeneous or heterogeneous schools (Frankenberg & Orfield, 2007), for example, and more research about what aspects of diversity provide children with positive experiences that contribute to promoting moral judgment and reasoning in development would be enlightening.

This chapter outlined a conceptual framework for research, summarized and systematized what we already know, and pointed to some of the exciting areas of inquiry about the development of moral reasoning. On the basis of the research reviewed here, we venture to propose that moral reasoning is essential to understanding the origins and development of human morality.

REFERENCES

Abrams, D., & Rutland, A. (2008). The development of subjective group dynamics. In S. Levy & M. Killen (Eds.), *Intergroup Relations and Attitudes in Childhood through Adulthood* (pp. 47–65). Oxford, UK: Oxford University Press.

Abrams, D., Rutland, A., Pelletier, J. G., & Ferrell, J. M. (2009). Children's group nous: understanding and applying peer exclusion within and between groups. *Child Development*, *80*, 224–243. doi:10.1111/j.1467-8624.2008.01256.x

Adler, J. E. (2008). Introduction: philosophical foundations. In J. E. Adler & L. J. Rips (Eds.), *Reasoning: Studies of Human Inference and Its Foundations* (pp. 1–34). Cambridge, UK: Cambridge University Press.

Appiah, K. A. (2005). *The Ethics of Identity*. Princeton, NJ: Princeton University Press.

Ball, C. L., Smetana, J. G., & Sturge-Apple, M. L. (2017). Following my head and my heart: integrating preschoolers' empathy, theory of mind, and moral judgments. *Child Development*, *88*, 597–611. https://doi.org/10.1111/cdev.12605

Barrett, K. C. (2005). The origins of social emotions and self-regulation in toddlerhood: New evidence. *Cognition and Emotion*, *19*, 953–979. doi:10.1080/02699930500172515

Barrett, K. C., Zahn-Waxler, C., & Cole, P. M. (1993). Avoiders vs. amenders: Implications for the investigation of guilt and shame during toddlerhood? *Cognition and Emotion*, *7*, 481–505. doi:10.1080/02699939308409201

Bennett, M., & Sani, F. (2004). *The Development of the Social Self*. New York, NY: Psychology Press. doi:10.4324/9780203391099

Blake, P. R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*, 215–224. doi:10.1016/j.cognition.2011.04.006

Bretherton, I., & Bates, E. (1979). The emergence of intentional communication. *New Directions for Child and Adolescent Development*, *1979*, 81–100. doi:10.1002/cd.23219790407

Brown, C. S., & Bigler, R. S. (2005). Children's perceptions of discrimination: A developmental model. *Child Development*, *76*, 533–553.

Brown, R. J., & Gaertner, S. L. (2001). *Blackwell Handbook in Social Psychology*. Oxford, UK: Blackwell.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*, 55–72. doi:10.1111/j.2044–835X.1991.tb00862.x

Campos, J. J., Witherington, D., Anderson, D. I., Frankel, C. I., Uchiyama, I., & Barbu-Roth, M. (2008). Rediscovering development in infancy. *Child Development*, *79*, 1625–1632. doi:10.1111/j.1467-8624.2008.01212.x

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*, 1–174. doi:10.2307/1166214

Cohn, J. F., & Tronick, E. Z. (1983). Three-month-old infants' reaction to simulated maternal depression. *Child Development*, *54*, 185–193. doi:10.2307/1129876

Cole, P. M., Barrett, K. C., & Zahn-Waxler, C. (1992). Emotion displays in two-year-olds during mishaps. *Child Development*, *63*, 314–324. doi:10.1111/j.1467-8624.1992.tb01629.x

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. *Psychological Science*, *17*, 1082–1089.

Dahl, A. (2014). Definitions and developmental processes in research on infant morality. *Human Development*, *57*, 241–249. doi:10.1159/000364919

Dahl, A. (2015). The developing social context of infant helping in two U.S. samples. *Child Development*, *86*, 1080–1093. doi:10.1111/cdev.12361

Dahl, A. (2016a). Infants' unprovoked acts of force toward others. *Developmental Science*, *19*, 1049–1057. https://doi.org/10.1111/desc.12342

Dahl, A. (2016b). Mothers' insistence when prohibiting infants from harming others in everyday interactions. *Frontiers in Psychology*, *1448*. https://doi.org/10.3389/fpsyg.2016.01448

Dahl, A. (2017). Ecological commitments: Why developmental science needs naturalistic methods. *Child Development Perspectives*, *11*, 79–84. doi:10.1111/cdep.12217

Dahl, A., & Campos, J. J. (2013). Domain differences in early social interactions. *Child Development*, *84*, 817–825. doi:10.1111/cdev.12002

Dahl, A., & Freda, G. F. (2017). How young children come to view harming others as wrong: A developmental analysis. In J. A. Sommerville & J. Decety (Eds.), *Social Cognition* (pp. 151–184). New York, NY: Routledge.

Dahl, A., & Kim, L. (2014). Why is it bad to make a mess? Preschoolers' conceptions of pragmatic norms. *Cognitive Development*, *32*, 12–22. doi:10.1016/j.cogdev.2014.05.004

Dahl, A., Satlof-Bedrick, E., Hammond, S. I., Drummond, J. D., Waugh, W. E., & Brownell, C. A. (in press). Explicit scaffolding increases helping in younger infants. *Developmental Psychology*, *53*, 407–416. doi:10.1037/dev0000244

Dahl, A., Schuck, R. K., & Campos, J. J. (2013). Do young toddlers act on their social preferences? *Developmental Psychology*, *49*, 1964–1970. doi:10.1037/a0031460

Dahl, A., Sherlock, B. R., Campos, J. J., & Theunissen, F. E. (2014). Mothers' tone of voice depends on the nature of infants' transgressions. *Emotion*, *14*, 651–665. doi:10.1037/a0036608

Dahl, A., Uttich, K., Gingo, M., & Turiel, E. (2013, June). "People can reason about moral dilemmas (when researchers allow them to)." Paper presented at the annual meeting of the Jean Piaget Society, Chicago, IL.

Damon, W. (1977). *The Social World of the Child*. San Francisco, CA: Jossey-Bass.

Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., & Knafo, A. (2013). Concern for others in the first year of life: Theory, evidence, and avenues for research. *Child Development Perspectives*, *7*, 126–131. doi:10.1111/cdep.12028

Drummond, J. D., Hammond, S. I., Satlof-Bedrick, E., Waugh, W. E., & Brownell, C. A. (2017). Helping the one you hurt: Toddlers' rudimentary guilt, shame, and prosocial behavior after harming another. *Child Development*, *88*, 1382–1397. https://doi.org/10.1111/cdev.12653

Dunn, J. (1988). *The Beginnings of Social Understanding*. Cambridge, MA: Harvard University Press.

Dunn, J., & Munn, P. (1985). Becoming a family member: family conflict and the development of social understanding in the second year. *Child Development*, *56*, 480–492. doi:10.2307/1129735

Dunn, J., & Munn, P. (1986). Siblings and the development of prosocial behaviour. *International Journal of Behavioral Development*, *9*, 265–284.

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, *51*, 665–697.

Eisenberg, N., Spinrad, T. L., & Knafo-Noam, A. (2015). Prosocial development. In M. E. Lamb & R. M. Lerner (Eds.), *Handbook of Child Psychology and Developmental Science* (7th ed., Vol. 3, pp. 610–656). Hoboken, NJ: Wiley.

Elenbaas, L., & Killen, M. (2016). Research in developmental psychology: social exclusion among children and adolescents. In P. Riva & J. Eck (Eds.), *Social Exclusion: Psychological Approaches to Understanding and Reducing Its Impact* (pp.89–108). New York, NY: Springer. doi:10.1007/978-3-319-33033-4

Elenbaas, L., Rizzo, M.T., Cooley, S., & Killen, M. (2016). Rectifying or perpetuating resource disparities: Children's responses to social inequalities based on race. *Cognition*, *155*, 176–187. https://doi.org/10.1016/j.cognition.2016.07.002

Elqayam, S., & Evans, J. S. B. T. (2011). Subtracting "ought" from "is": Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, *34*, 233–248. https://doi.org/10.1017/S0140525X1100001X

Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford Review*, *5*, 5–15.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998. https://doi.org/10.1126/science.1218633

Frankenberg, E., & Orfield, G. (Eds.). (2007). *Lessons in Integration: Realizing the Promise of Racial Diversity in American Schools*. Charlottesville, VA: University of Virginia Press.

Gaertner, S. L., & Dovidio, J. F. (2005) Categorization, recategorization, and intergroup bias. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the Nature of Prejudice: Fifty Years after Allport* (pp. 71–88). Oxford, UK: Blackwell. doi:10.1002/9780470773963.ch5

Gewirth, A. (1978). *Reason and Morality*. Chicago, IL: University of Chicago Press.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101–124. doi:10.1080/1047840X.2012.651387

Greene, J. D. (2014). Beyond point-and-shoot morality: Why cognitive (neuro)science matters for ethics. *Ethics*, *124*, 695–726. doi:10.1086/675875

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834. doi:10.1037/0033-295X.108.4.814

Haidt, J. (2008). Morality. *Perspectives on Psychological Science*, *3*, 65–72.

Haidt, J., & Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral Psychology. Vol 2: The Cognitive Science of Morality: Intuition and Diversity* (pp. 181–217). Cambridge, MA: MIT Press.

Haidt, J., Bjorklund, F., & Murphy, S. (2000). "Moral dumbfounding: when intuition finds no reason." Unpublished manuscript. Retrieved from http://www.faculty.virginia.edu/haidtlab/articles/manuscripts/haidt.bjorklund.working-paper.when%20intuition%20finds%20no%20reason.pub603.doc

Hamlin, J. K. (2014). The origins of human morality: Complex socio-moral evaluations by preverbal infants. In J. Decety & Y. Christen (Eds.), *New Frontiers in Social Neuroscience* (Vol. 21, pp. 165–188). Cham, Switzerland: Springer International.

Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not like me = bad: infants prefer those who harm dissimilar others. *Psychological Science*, *24*, 589–594. doi:10.1177/0956797612457785

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. doi:10.1038/nature06288

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*, 30–39.

Hammond, S. I., & Carpendale, J. I. M. (2015). Helping children help: The relation between maternal scaffolding and children's early help *Social Development*, *24*, 367–383. doi:10.1111/sode.12104

Harman, G. (1986). *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press.

Hastings, P. D., Miller, J. G., Kahle, S., & Zahn-Waxler, C. (2014). The neurobiological bases of empathic concern for others. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (2nd ed., pp. 411–434). New York, NY: Psychology Press.

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & Language*, *22*, 1–21.

Hay, D. F. (2005). The beginnings of aggression in infancy. In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Developmental Origins of Aggression* (pp. 107–132). New York, NY: Guilford Press.

Hay, D. F., Nash, A., & Pedersen, J. (1981). Responses of six-month-olds to the distress of their peers. *Child Development*, *52*, 1071–1075. doi:10.2307/1129114

Helwig, C. C., Ruck, M. D., & Peterson-Badali, M. (2014). Rights, civil liberties, and democracy. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (2nd ed., pp. 46–69). New York, NY: Psychology Press.

Hepach, R., Vaish, A., & Tomasello, M. (2016). Children's intrinsic motivation to provide help themselves after accidentally harming others. *Child Development*, *88*(4), 1251–1264. https://doi.org/10.1111/cdev.12646

Hitti, A., Mulvey, K. L., Rutland, A., Abrams, D., & Killen, M. A. (2013). When is it okay to exclude a member of the ingroup? Children's and adolescents' social reasoning. *Social Development*, *23*, 451–469. doi: 10.1111/sode.12047

Hoffman, M. L. (2000). *Empathy and Moral Development: Implications for Caring and Justice*. New York, NY: Cambridge University Press.

Horn, S. S. (2012). Contemporary attitudes about sexual orientation. In C. Patterson & A. D'Augelli (Eds.), *Handbook for the Psychology of Sexual Orientation* (pp. 239–251). Oxford, UK: Oxford University Press.

Jacobson, D. (2012). Moral dumbfounding and moral stupefaction. In M. Timmons (Ed.), *Oxford Studies in Normative Ethics* (Vol. 2, pp. 289–316). Oxford, UK: Oxford University Press.

Kagan, J. (1981). *The Second Year: The Emergence of Self-Awareness*. Cambridge, MA: Harvard University Press.

Kaler, S. R., & Kopp, C. B. (1990). Compliance and comprehension in very young toddlers. *Child Development*, *61*, 1997–2003. doi:10.2307/1130853

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. *PloS One*, *7*, e43979. doi:10.1371/journal.pone.0043979

Kant, I. (1785/1959). *Foundations of the Metaphysics of Morals*. New York, NY: Bobbs-Merrill.

Kayyal, M. H., Pochedly, J., McCarthy, A., & Russell, J. A. (2015). On the limits of the relation of disgust to judgments of immorality. *Frontiers in Psychology*, *6*. doi:10.3389/fpsyg.2015.00951

Kenward, B., & Dahl, M. (2011). Preschoolers distribute scarce resources according to the moral valence of recipients' previous actions. *Developmental Psychology*, *47*, 1054–1064. doi:0.1037/a0023869 10.1037/a0023869.supp

Kihlstrom, J. F. (2013). Threats to reason in moral judgment. *Hedgehog Review*, *15*, 8–18.

Killen, M., Elenbaas, L., & Rutland, A. (2015). Balancing the fair treatment of others while preserving group identity and autonomy. *Human Development*, *58*, 253–272. doi:10.1159/000444151

Killen, M., Mulvey, K. L., & Hitti, A. (2013). Social exclusion in childhood: A developmental intergroup perspective. *Child Development*, *84*, 772–790. doi:10.1111/cdev.12012

Killen, M., Mulvey, K. L., Richardson, C. B., Jampol, N., & Woodward, A. (2011). The "accidental transgressor": Morally-relevant theory of mind. *Cognition*, *119*, 197–215. doi:10.1016/j.cognition.2011.01.006

Killen, M., Pisacane, K., Lee-Kim, J., & Ardila-Rey, A. (2001). Fairness or stereotypes?

Young children's priorities when evaluating group exclusion and inclusion. *Developmental Psychology*, *37*, 587–596. doi:10.1037/0012-1649.37.5.587

Killen, M., & Rutland, A. (2011). *Children and Social Exclusion: Morality, Prejudice, and Group Identity*. Malden, MA: Wiley-Blackwell. doi:10.1002/9781444396317

Killen, M., Rutland, A., Abrams, D., Mulvey, K. L., & Hitti, A. (2013). Development of intra- and intergroup judgments in the context of moral and social-conventional norms. *Child Development*, *84*, 1063–1080. doi:10.1111/cdev.12011

Killen, M., & Smetana, J. G. (1999). Social interactions in preschool classrooms and the development of young children's conceptions of the personal. *Child Development*, *70*, 486–501. doi:10.1111/1467-8624.00035

Killen, M., & Smetana, J. G. (2015). Origins and development of morality. In R. M. Lerner & M. Lamb (Eds.), *Handbook of Child Psychology and Developmental Science* (7th ed., Vol. 3, pp. 701–749). New York, NY: Wiley-Blackwell.

Klimes-Dougan, B., & Kopp, C. B. (1999). Children's conflict tactics with mothers: A longitudinal investigation of the toddler and preschool years. *Merrill-Palmer Quarterly*, *45*, 226–241.

Kochanska, G., Gross, J. N., Lin, M.-H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, *73*, 461–482.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of Socialization Theory and Research* (pp. 347–480). Chicago, IL: Rand McNally.

Kuczynski, L., Kochanska, G., Radke-Yarrow, M., & Girnius-Brown, O. (1987). A developmental interpretation of young children's noncompliance. *Developmental Psychology*, *23*, 799–806.

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, *10*, 518–536.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335–1342. https://doi.org/10.1126/science.208.4450.1335

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: conceptualizing, measuring and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, *67*, 1–183.

Martin, G. B., & Clark, R. D. (1982). Distress crying in neonates: species and peer specificity. *Developmental Psychology*, *18*, 3–9.

Mascolo, M. E., & Fischer, K. W. (2007). The co-development of self and sociomoral emotions during the toddler years. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional Development in the Toddler Years: Transitions and Transformations* (pp. 66–99). New York, NY: Guilford Press.

McGillicuddy-De Lisi, A., Daly, M., & Neal, A. (2006). Children's distributive justice judgments: Aversive racism in Euro-American children? *Child Development*, *77*, 1063–1080. doi:10.1111/j.1467-8624.2006.00919.x

Miller, P. A., Eisenberg, N., Fabes, R. A., & Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Developmental Psychology*, *32*, 210–219.

Mulvey, K. L. (2016). Children's reasoning about social exclusion: balancing many factors. *Child Development Perspectives*, *1*, 22–27. doi:10.1111/cdep.12157

Mulvey, K. L., Hitti, A., & Killen, M. (2013). Intentionality, morality, and exclusion: How children navigate the social world. In M. Banaji & S. Gelman (Eds.), *Navigating the Social World: What Infants, Children and Other Species Can Teach Us* (pp. 377–384). New York, NY: Oxford University Press.

Mulvey, K. L., Hitti, A., Rutland, A., Abrams, D., & Killen, M. (2014). Context differences in children's ingroup preferences. *Developmental Psychology*, *50*, 1507–1519. doi:10.1037/a0035593

Mulvey, K. L., & Killen, M. (2015). Challenging gender stereotypes: Resistance and exclusion.

Child Development, *86*, 681–694. doi:10.1111/cdev.12317

Nesdale, D., Durkin, K., Maass, A., & Griffiths, J. (2005). Threat, group identification, and children's ethnic prejudice. *Social Development*, *14*, 189–205. doi:10.1111/j.1467-9507.2005.00298.x

Nesdale, D., & Flesser, D. (2001). Social identity and the development of children's group attitudes. *Child Development*, *72*, 506–517. doi:10.1111/1467-8624.00293

Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development*, *49*, 400–407. doi:10.2307/1128704

Nucci, L., & Turiel, E. (2009). Capturing the complexity of moral development and education. *Mind, Brain, and Education*, *3*(3), 151–159.

Nucci, L. P., & Weber, E. K. (1995). Social interactions in the home and the development of young children's conceptions of the personal. *Child Development*, *66*, 1438–1452.

Nussbaum, M. C. (1999). *Sex and Social Justice*. Oxford, UK: Oxford University Press.

Oaksford, M., & Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in Cognitive Sciences*, *5*(8), 349–357.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in Cognitive Science*, *2*(3), 511–527. https://doi.org/10.1111/j.1756-8765.2010.01096.x

Payne, B. K., Jacoby, L. L., & Lambert, A. J. (2005). Attitudes as accessibility bias: Dissociating automatic and controlled processes. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The New Unconscious* (pp. 393–420). New York, NY: Oxford University Press.

Pettygrove, D. M., Hammond, S. I., Karahuta, E. L., Waugh, W. E., & Brownell, C. A. (2013). From cleaning up to helping out: Parental socialization and children's early prosocial behavior. *Infant Behavior and Development*, *36*, 843–846. doi:10.1016/j.infbeh.2013.09.005

Piaget, J. (1932). *The Moral Judgment of the Child*. New York, NY: Free Press.

Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: A comment on Haidt (2001). *Psychological Review*, *110*, 193–196. https://doi.org/10.1037/0033-295X.110.1.193

Pizarro, D., Inbar, Y., & Helion, C. (2011). On disgust and moral judgment. *Emotion Review*, *3*, 267–268. doi:10.1177/1754073911402394

Pons, F., Harris, P. L., & Rosnay, M. de. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*, 127–152. doi:10.1080/17405620344000022

Powell, L. J., & Spelke, E. S. (2014). "Third party preferences for imitation in preverbal infants." Paper presented at the International Conference on Infant Studies, Berlin, Germany.

Power, T. G., & Parke, R. D. (1986). Patterns of early socialization: Mother- and father-infant interaction in the home. *International Journal of Behavioral Development*, *9*, 331–341. doi:10.1177/016502548600900305

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental Psychology*, *44*, 875–881. doi:10.1037/0012–1649.44.3.875

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*, 12–21.

Ridgeway, C. L. (2013). Why status matters for inequality. *American Sociological Review*, *79*, 1–16. doi:10.1177/0003122413515997

Rizzo, M. T., Elenbaas, L., Cooley, S., & Killen, M. (2016). Children's recognition of fairness and others' welfare in a resource allocation task: Age-related changes. *Developmental Psychology*, *52*, 1307–1317. http://dx.doi.org/10.1037/dev0000134

Rizzo, M. T., & Killen, M. (2016). Children's understanding of equity in the context of inequality. *British Journal of Developmental Psychology*, *34*, 569–581. doi:10.1111/bjdp.12150

Rogoff, B. (2003). *The Cultural Nature of Human Development*. Oxford, UK: Oxford University Press.

Rossano, F., Rakoczy, H., & Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition*, *121*, 219–227. doi:10.1016/j.cognition.2011.06.007

Roth-Hanania, R., Davidov, M., & Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, *34*, 447–458. doi:10.1016/j.infbeh.2011.04.007

Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment & Decision Making*, *10*, 296–313.

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of Child Psychology: Social, Emotional, and Personality Development* (6th ed., Vol. 3, pp. 571–645). Hoboken, NJ: Wiley.

Ruble, D. N., Martin, C. L., & Berenbaum, S. (2006). Gender development. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology: Personality and Social Development* (Vol. 3, 6th ed., pp. 858–932). Hoboken, NJ: Wiley.

Rutland, A., & Killen, M. (2015). A developmental science approach to reducing prejudice and social exclusion: Intergroup processes, social-cognitive development, and moral reasoning. *Social Issues and Policy Review*, *9*, 121–154.

Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Science*, *5*, 279–291. doi:10.1177/1745691610369468

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology*, *12*, 175–176. doi:10.1037/0012-1649.12.2.175

Scanlon, T. M. (2014). *Being Realistic about Reasons*. Oxford, UK: Oxford University Press.

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2013). Young children understand and defend the entitlements of others. *Journal of Experimental Child Psychology*, *116*, 930–944. http://doi.org/10.1016/j.jecp.2013.06.013

Schmidt, M. F. H., Svetlova, M., Johe, J., & Tomasello, M. (2016). Children's developing understanding of legitimate reasons for allocating resources unequally. *Cognitive Development*, *37*, 42–52. http://doi.org/10.1016/j.cogdev.2015.11.001

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, *34*, 1096–1109. doi:10.1177/0146167208317771

Searle, J. R. (2003). *Rationality in Action*. Cambridge, MA: MIT Press.

Sen, A. K. (2009). *The Idea of Justice*. Cambridge, MA: Harvard University Press.

Shaw, A., & Olson, K. R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Child Psychology*, *141*, 382–395. doi:10.1037/a0025907

Siegal, M., & Storey, R. M. (1985). Day care and children's conceptions of moral and social rules. *Child Development*, *56*, 1001–1008. doi:10.2307/1130111

Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology*, *25*, 499–508.

Smetana, J. G. (2011). *Adolescents, Families, and Social Development: How Teens Construct Their Worlds*. New York, NY: Wiley/Blackwell.

Smetana, J. G. (2013). Moral development: The social domain theory view. In P. D. Zelazo (Ed.), *Oxford Handbook of Developmental Psychology* (Vol. 1, pp. 832–866). Oxford, UK: Oxford University Press.

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly*, *36*, 329–346.

Smetana, J. G., Jambon, M., & Ball, C. (2014). The social domain approach to children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (2nd ed., pp. 23–45). New York, NY: Psychology Press.

Smetana, J. G., Rote, W. M., Jambon, M., Tasopoulos-Chan, M., Villalobos, M., & Comer, J. (2012). Developmental changes and individual differences in young children's moral judgments. *Child Development*, 683–696. doi:10.1111/j.1467-8624.2011.01714.x

Sodian, B., Licata, M., Kristen, S., Paulus, M., Killen, M., & Woodward, A. L. (2016). Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: A longitudinal investigation. *Child Development*, *87*, 1221–1232. doi: 10.1111/cdev.12533

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*, B1–B11. doi:10.1016/j.cognition.2004.07.004

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, *21*, 195–200. doi:10.1037/0012–1649.21.1.195

Sroufe, L. A. (1996). *Emotional Development: The Organization of Emotional Life in the Early Years*. New York, NY: Cambridge University Press.

Stipek, D. J., Gralinski, J. H., & Kopp, C. B. (1990). Self-concept development in the toddler years. *Developmental Psychology*, *26*, 972–977.

Stipek, D. J., Recchia, S., McClintic, S., & Lewis, M. (1992). Self-evaluation in young children. *Monographs of the Society for Research in Child Development*, *57*, 100. doi:10.2307/1166190

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child Development*, *81*, 1814–1827.

Thijs, J., Verkuyten, M., & Grundel, M. (2014). Ethnic classroom composition and peer victimization: the moderating role of classroom attitudes. *Journal of Social Issues*, *70*, 134–150.

Thompson, R. A., & Goodvin, R. (2007). Taming the tempest in the teapot: Emotion regulation in toddlers. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional Development in the Toddler Years: Transitions and Transformations* (pp. 320–341). New York, NY: Guilford Press.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *Monist*, *59*, 204–217.

Tisak, M. S. (1993). Preschool children's judgments of moral and personal events involving physical harm and property damage. *Merrill-Palmer Quarterly*, *39*, 375–390.

Tisak, M. S., & Turiel, E. (1988). Variation in seriousness of transgressions and children's moral and conventional concepts. *Developmental Psychology*, *24*, 352–357.

Tomasello, M. (2016). *A Natural History of Human Morality*. Cambridge, MA: Harvard University Press.

Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech: The Beginning of Human Communication* (pp. 321–348). Cambridge, UK: Cambridge University Press.

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist*, *44*, 112–119. doi:10.1037/0003-066X.44.2.112

Tropp, L. R., & Prenovost, M. A. (2008). The role of intergroup contact in predicting children's inter-ethnic attitudes: Evidence from meta-analytic and field studies. In S. Levy & M. Killen (Eds.), *Intergroup Attitudes and Relations in Childhood through Adulthood* (pp. 236–248). Oxford, UK: Oxford University Press.

Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge, UK: Cambridge University Press.

Turiel, E. (2002). *The Culture of Morality: Social Development, Context, and Conflict*. Cambridge, UK: Cambridge University Press.

Turiel, E., (2008). Thoughts about actions in social domains: Morality, social conventions, and social interactions. *Cognitive Development*, *23*, 136–154.

Turiel, E. (2014). Morality: Epistemology, development, and social opposition. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (2nd ed., pp. 3–22). New York, NY: Psychology Press.

Turiel, E., & Dahl, A. (in press). The development of domains of moral and conventional norms, coordination in decision-making, and the implications of social opposition. In K. Bayertz & N. Roughley (Eds.), *The Normative Animal: On the Biological Significance of Social, Moral,* *and Linguistic Norms*. Oxford, UK: Oxford University Press.

Turiel, E., Hildebrandt, C., & Wainryb, C. (1991). Judging social issues: Difficulties, inconsistencies, and consistencies. *Monographs of the Society for Research in Child Development*, *56*, 1–116. doi:10.2307/1166056

Turiel, E., & Killen, M. (2010). Taking emotions seriously: the role of emotions in moral development. In W. Arsenio & E. Lemerise (Eds.), *Emotions in Aggression and Moral Development* (pp. 33–52). Washington, DC: American Psychological Association.

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development*, *81*, 1661–1669.

Vaish, A., Carpenter, M., & Tomasello, M. (2016). The early emergence of guilt-motivated prosocial behavior. *Child Development*, *87*, 1772–1782. https://doi.org/10.1111/cdev.12628

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions: Children intervene in moral transgressions. *British Journal of Developmental Psychology*, *29*, 124–130. doi:10.1348/026151010X532888

Verkuyten, M., & Thijs, J. (2006). Ethnic discrimination and global self-worth in early adolescents: The mediating role of ethnic self-esteem. *International Journal of Behavioral Development*, *30*, 107–116.

Wainryb, C. (1991). Understanding differences in moral judgments: The role of informational assumptions. *Child Development*, *62*, 840–851. doi:10.2307/1131181

Wainryb, C., Brehl, B. A., & Matwin, S. (2005). Being hurt and hurting others: Children's narrative accounts and moral judgments of their own interpersonal conflicts. *Monographs of the Society for Research in Child Development*, *70*. https://doi.org/10.1111/j.1540–5834.2005.00350.x

Wainryb, C., & Pasupathi, M. (2009). Developing moral agency in the midst of violence: children, political conflict, and values. In W. McCormack (Ed.), *Values and Violence: Intangible Aspects of Terrorism* (Vol. 4, pp. 169–187). New York, NY: Springer. doi:10.1007/978-1-4020-8660-1_11

Wainryb, C., & Turiel, E. (1994). Dominance, subordination, and concepts of personal entitlements in cultural contexts. *Child Development*, *65*(6), 1701–1722. https://doi.org/10.2307/1131289

Warneken, F. (2013). Young children proactively remedy unnoticed accidents. *Cognition*, *126*, 101–108. doi:10.1016/j.cognition.2012.09.011

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303. doi:10.1126/science.1121448

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*, 271–294.

Warneken, F., & Tomasello, M. (2013). Parental presence and encouragement do not influence helping in young children: Parental influence on helping in young children. *Infancy*, *18*, 345–368. doi:10.1111/j.1532-7078.2012.00120.x

Waugh, W. E., Brownell, C. A., & Pollock, B. (2015). Early socialization of prosocial behavior: patterns in parents' encouragement of toddlers' helping in an everyday household task. *Infant Behavior and Development*, *39*, 1–10.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*, 502–517.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*, 780–784.

Willems, E. P. (1967). Toward an explicit rationale for naturalistic research methods. *Human Development*, *10*, 138–154. doi:10.1159/000270581

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34. doi:10.1016/S0010-0277 00058-4

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, *18*, 53–57. doi:10.1111/j.1467-8721.2009.01605.x

Zahn-Waxler, C., & Kochanska, G. (1990). The origins of guilt. In R. A. Thompson (Ed.), *Nebraska Symposium on Motivation* (Vol. 36, pp. 183–258). Lincoln, NE: University of Nebraska Press.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28*, 126–136. doi:10.1037/0012-1649.28.1.126

Young, S. K., Fox, N. A., & Zahn-Waxler, C. (1999). The relations between temperament and empathy in 2-year-olds. *Developmental Psychology*, *35*(5), 1189–1197. https://doi.org.oca.ucsc.edu/10.1037/0012-1649.35.5.1189

SOCIAL PSYCHOLOGY

CHAPTER 11

Attitudes

STEVEN T. BENGAL, JEREMY D. GRETTON, AND DUANE T. WEGENER

INTRODUCTION

From presidential debates to Super Bowl advertisements, from lobbyists to food critics, from the Gallup poll to friends on social media, from nationalism to prejudice, a tremendous amount of time, money, and effort is expended daily to understand or influence public opinion. Ubiquitous advertisements, in the mail, on the phone and radio, and in the form of banners on websites and e-mail spam, urge people to engage in particular actions—to save the environment, to switch cell phone service providers, to vote for a particular political candidate, or to eat, drink, or travel. In some places, stadiums shake with the roar of crowds cheering for their favorite team. In others, the roar heralds the toppling of governments. Although these examples initially might seem disconnected, they all relate directly to the attitudes people hold. Those attitudes might take the form of support for (or opposition of) a team or candidate. They might be (dis)liking of a person or product, or they might reflect a more favorable opinion of one's group than of another group. These many terms—attitudes, liking, support, favor, opinion—all refer to one's evaluation of people, ideas, or objects in one's environment. Thus, research on attitudes has addressed opinions (and related behaviors) related to everything from advertisements to education, healthcare to politics, sustainable energy to prejudice, and everywhere in between.

Attitudes: Definition and Related Constructs

An attitude is a summary evaluation of an object (i.e., a person, place, thing, or concept), ranging on a valenced mental dimension from good to bad, likable to dislikable, positive to negative (see Eagly & Chaiken, 1993; Petty, Wegener, & Fabrigar, 1997). An attitude thus may be thought of as a relatively positive or negative predisposition toward some entity, and often it also involves engaging with the attitude object behaviorally by approaching or avoiding (Krosnick, Judd, & Wittenbrink, 2005). These evaluations often are activated automatically and spontaneously upon encountering the attitude object (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996; Fazio, 1993). Even so, researchers have identified individual differences in people's need to evaluate (Jarvis & Petty, 1996) when there are not strong situational pressures toward doing so.

Early research into the structure of attitudes utilized a tripartite model (Rosenberg & Hovland, 1960) in which attitudes included affective, behavioral, and cognitive components (or, alternatively, affect, beliefs, and conation; for related research, see Breckler, 1984; Ostrom, 1969). More recent definitions and conceptualizations, however, have limited the term "attitude" to refer to the summary evaluation, with affect, cognition, and behavior as potential antecedents to the attitude and as potential consequences of the attitude (see Petty & Wegener, 1998). Defining "attitude" in this way has led to examination of differences between attitudes based primarily on affect or cognition (e.g., Fabrigar & Petty, 1999) or attitudes for which there is evaluative inconsistency (ambivalence) across the components underlying the evaluation (see Fabrigar, MacDonald, & Wegener, 2005).

The related concept of affect has been (and continues to be) used in different ways that can create confusion. For example, researchers have used the term "affect" to refer to liking or summary evaluation of an object (Zajonc, 1980). For example, in work on mere exposure effects, the "affective discriminations" were evaluative preferences (e.g., Kunst-Wilson & Zajonc, 1980). Similarly, work on the affect heuristic (Slovic, Finucane, Peters, & MacGregor, 2002) addressed use of liking of choice options prior to the choice as a "shortcut" to decision making. Such liking could be based on affective/emotional reactions to the options, but need not be. For clarity, attitudes researchers have come to restrict the term "affect" to feeling states, such as moods and emotions. When an attitude object elicits an emotional reaction or an incidental mood state supplies feelings that are misattributed to the object (Schwarz, 1990), the attitude can be based on affect. These affective influences can be powerful, and some have suggested a sort of "primacy" for affect in attitude formation, above and beyond the other two components (see Cervellon & Dube, 2002).

In the tripartite days of attitudes research, the cognitive component referred specifically to thoughts or beliefs about the attitude object. A key distinction between beliefs and attitudes more broadly is that a belief need not have immediate evaluative implications, whereas an attitude is defined as the evaluation itself. For example, a person could believe that a target individual is very tall. Yet tallness might have positive evaluative implications if one is seeking to identify a potential basketball center but have negative evaluative implications if one is seeking to identify a potential horseracing jockey. In recent years, at least within social psychology, the term "cognition" has broadened beyond (relatively explicit) thoughts or beliefs to address a wide variety of processes that involve the brain (see Wegener & Carlston, 2005). This broadening of the concept of cognition might make it more difficult to distinguish cognitive from noncognitive processes, but it reflects the increasing emphasis on brain mechanisms in all manner of psychological action and reaction. As we discuss later in the chapter, in areas of attitude functions, attitude strength, and particularly implicit attitude measurement techniques, postulated roles for memory in attitude formation, change, and persistence over time also reflect key cognitive concepts related to attitudes.

Behavior has been studied both as influencing attitudes (e.g., through cognitive dissonance, e.g., Festinger, 1957, or self-perception, e.g., Bem, 1972; see the section titled "Persuasion: Attitude Change" and Albarracín & Wyer, 2000) and as following from attitudes (e.g., Glasman & Albarracín, 2006). Just as attitudes relate to the valence dimension, behavioral activity often is arrayed from approach (the positive) to avoidance (the negative). Though not originally discussed as such, neuroscience has provided some evidence that attitudes may activate specific regions in the motor cortex of the brain, as a sort of preparation for action (see Preston & de Waal, 2002).

Measurement: Direct and Indirect

Many descriptions of attitudes have followed from the measurement tools being used at the time, and advancements in attitude theory often co-occur with new measurement tools (Ostrom, 1989). Common measures might vary from domain to domain, but in the psychological literature, attitude measures can largely be discussed as being either direct (explicit) or indirect (implicit) measures (see Krosnick et al., 2005). The theoretical assumptions differ across direct versus indirect measures, and it is important for researchers to align the qualities of the measure to the research questions of interest (see Wegener & Fabrigar, 2004).

Direct Measures

In many domains, such as consumer or (often) political preferences, it seems reasonable to assume that respondents are willing and able to share their opinions. In such cases, researchers or pollsters typically ask directly for respondents' attitudes. Specific attitude measures vary, but the direct measures generally involve a form of overt self-report on a valenced scale of favorability toward the object of interest.

In some of the earliest work on attitude measurement, Thurstone (1928) adapted paired-comparison methods from psychophysics (e.g., relating stimulus qualities, such as decibels, to perceptions of loudness) to index levels of favorability of people's attitudes. Because these methods were cumbersome (e.g., they required stimulus scaling on a separate set of respondents from those using the scale to report attitudes), researchers soon developed more manageable methods. An early option that continues to be popular was Likert scaling (Likert, 1932). A Likert scale consists of statements either clearly favoring or clearly opposing a particular object or position, and respondents are asked to rate their amount of agreement or disagreement with each statement. Likert scales measure (dis)agreement with a set of statements and determine which statements best hang together using item-total correlations. Overall attitudes are based on those items with sufficient item-total correlations and are indexed by the mean agreement with statements favoring the object (and disagreement with statements opposing the object).

Even Likert scales require different statements for different attitude objects. In contrast, the work of Osgood, Suci, and Tannenbaum (1957) on semantic differential scales identified a set of bipolar adjectives that were broadly applicable to almost any attitude object. Semantic differential scales involve rating an attitude object on a scale anchored by bipolar evaluative adjectives (e.g., good–bad, positive–negative, beneficial–harmful, wise–foolish, pleasant–unpleasant, etc.). In other words, the semantic content (word meaning) falls along a differential (a difference gradient). Because many of the adjective pairs apply across a wide array of potential attitude objects, they are used in many settings. Researchers often examine inter-item correlations or conduct factor analyses to ensure that the items being used sufficiently hang together and likely assess a single construct. (For additional discussion, see Wegener & Fabrigar, 2004.)

Self-report scales are widely used. They are easy for researchers to design and for people to understand and answer. However, researchers acknowledge that the content of an attitude that comes to mind is separable from the rating that one gives in that the person must translate the attitude into a response that makes sense on that scale (e.g., Tourangeau & Rasinski, 1988). One must take great care in designing the format of these scales, as respondents may infer information about how to answer a question based on how the question is presented. Everything from the labeling of scale points to the order of response alternatives has been found to impact responses to self-report measures. For example, respondents can infer that the midpoint of a scale represents the "average" or mean response and can modify their responses as a consequence of this inference (Schwarz, 1999).

Another concern that has prompted continued development of new attitude measures is the issue of whether participants are willing to tell the truth on an attitude scale. People may screen their direct responses when audience attitudes are known or particular attitudes are viewed as more socially desirable. For example, a person might profess different political beliefs in a group of Republicans rather than Democrats. When using direct measures, one method of reducing self-presentation concerns is assuring anonymity of responses (see Michaelis & Eysenck, 1971). However, sometimes even assurances of anonymity may not fully address self-presentation concerns. In a somewhat notorious technique called the bogus pipeline, respondents' attitudes toward noncontroversial topics are gathered surreptitiously (e.g., as part of another study) and the respondent is first convinced by a "calibration phase" that a machine to which they are connected (such as facial electromyography or electroencephalography) can detect true or false responses. In a later phase involving the more controversial questions, the respondent has no incentive to lie (and sometimes is told that the questions are tests of whether the person, in fact, knows his or her own opinions or not). Use of the bogus pipeline has increased the reports of socially sensitive attitudes (such as racial prejudice; Sigall & Page, 1971) and has increased relations of self-report measures with indirect/implicit measures (e.g., Nier, 2005).

Beyond intentional deception by respondents, there is also the possibility that, at times, participants may not actually be able to provide their attitude accurately. Whether they are deceiving themselves about how they really feel or unintentionally misconstruing their position, personal deception cannot be resolved through any of the aforementioned self-report measures. In an attempt to address both potential distortion of self-reports and potential inability to report, a variety of indirect measures have been developed.

Indirect Measures

Indirect measures are aimed at inferring a person's attitude from some type of judgment or behavior without asking the person to report an opinion. In one early indirect measure, Hammond (1948) attempted to measure respondents' attitudes through a supposed test of facts in which none of the available answers was correct. In this error choice method, the respondent's attitude was inferred based on the types of errors chosen (e.g., support for U.S. military involvement in Afghanistan might be inferred from choosing overly low rather than high estimates of the number of civilian casualties caused by U.S. activities).

Other types of indirect measures also have been used, such as those using physical proximity as a proxy for evaluation. For example, in research on prejudice, a number of researchers have surreptitiously measured how closely research participants sit to a member of a stigmatized group who is already present (e.g., Westie, 1953) or whose location is implied by objects presumably belonging to the person (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994; see also Fazio, Jackson, Dunton, & Williams, 1995). Importantly, these incidental behaviors often are performed without conscious recognition of the evaluative signals they send, so participants are less likely to screen such responses.

Physiological indirect measures have an underlying assumption that people have different physical reactions to objects that they like rather than dislike. One critical issue with some of the physiological or neurological methods is that they tend to fail to distinguish between the positive or negative valence of an attitude and instead primarily detect a degree of physiological arousal or motivational relevance (see Petty & Cacioppo, 1996, for a review of older physiological measures, such as pupillary dilation or heart rate; see Cunningham, Van Bavel, & Johnsen, 2008, for a discussion of amygdala activation as also measuring relevance in addition to valence). However, a number of physiological measures do seem to differentiate successfully between positively and negatively valenced reactions. For instance, event-related potentials have been used to identify differences in evaluations between "baseline" stimuli and an "oddball" (surprising) target stimulus. In particular, the spike in electrical activity occurring approximately 300 ms after a category shift in the oddball paradigm has been used to index the amount of difference in valence between normatively positive or negative baseline words and a neutral or differently valenced oddball word. The size of the electrical spike in activity has proven successful at distinguishing both the valence of attitudes toward the oddball stimulus and the amount of discrepancy between attitudes toward the baseline stimuli and the oddball stimulus (even when respondents are deliberately attempting to be duplicitous; Crites, Cacioppo, Gardner, & Berntson, 1995). Another physiological measure indexes electrical activity in facial muscles using electromyography. When research participants listen to persuasive material they find agreeable, activity in the zygomatic (smiling) muscles increases, and when they listen to disagreeable material, activity in the corrugator (frowning) muscles increases (Cacioppo & Petty, 1979; although more recent research suggests that corrugator activity produces stronger linear effects of valence; Larsen, Norris, & Cacioppo, 2003).

A final and more recent physiological approach uses functional magnetic resonance imaging (fMRI) to observe blood flow in different brain areas as a function of the presentation of valenced psychological objects. Although initial work using this imaging produced greater activation in the amygdala for negative stimuli compared to positive stimuli (e.g., Zald & Pardo, 1997), later work also produced greater activation for positive than neutral stimuli (see Liberzon, Phan, Decker, & Taylor, 2003), suggesting that the amygdala may be involved with identification of more motivationally relevant stimuli (Cunningham et al., 2008). Recent research has shifted from locating isolated brain areas associated with attitudes to investigating patterns of activation across many brain regions, treating evaluation as more distributed and dynamic than in the earlier research (e.g., Cunningham, Zelazo, Packer, & Van Bavel, 2007; Van Bavel, Xiao, & Cunningham, 2012).

Because some of the earliest indirect measures had generally lower levels of reliability and (possibly) validity (Lemon, 1973), most of the early indirect measures are little used today (with the exception of some use of physical proximity behavioral measures). Also, because of the expense and time that goes into physiological and neurological measures, these indices have not been used frequently as indirect measures of attitudes. However, following developments in cognitive psychology and in social cognition, a number of indirect measures based on response time and response competition have gained popularity and continued to develop and push theory to this day. We discuss these types of measures later in the chapter in the section titled "Recent/Emerging Trends."

BACKGROUND ISSUES

Eras in Attitude Theory

Attitude is not a new concept in psychology. Indeed, almost a century ago, it was regarded as such an essential topic that authors such as Watson (1925) discussed the entire field of social psychology as the study of attitudes. Although much has changed in psychology in the last 100 years, the study of attitudes has persisted. Constant development of new methods and theories and the explanatory power that attitude models have for behavior have kept the attitudes domain vibrant throughout its long history. Commentators on the development of attitudes research have identified a number of distinct eras (e.g., Briñol & Petty, 2012).

Early work on attitudes focused on attitude measurement (e.g., Likert, 1932; Thurstone & Chave, 1929). The tools developed by these early researchers persist to the modern day and, especially in the case of Likert scales, still are commonly employed. Early attempts to validate the attitude measures also provoked interest in behavioral consequences of attitudes (e.g., see Newcomb, 1943).

The research emphasis on attitudes and persuasion dramatically increased following World War II. The Allies had faced active propaganda efforts by both the Nazis and the Japanese, and they sought to influence the European continent without a physical presence on the continent (e.g., through Radio Free Europe and other efforts). Psychologists, such as Leonard Doob and Carl Hovland, who were involved in Allied efforts during the war, realized how little was known about how attitudes were formed and changed, and they made a concerted effort to study such questions systematically both during the war (see Hovland, Lumsdaine, & Sheffield, 1949) and when they returned to Yale University after the war. Hovland is generally credited with drawing researchers to Yale throughout the 1950s to study communication and persuasion (e.g., see Hovland, Janis, & Kelley, 1953). In addition to the Yale Group efforts, the 1950s and 1960s saw the examination of many specific theories related to attitude formation and change. These included applications of classical and operant conditioning (e.g., Insko, 1965; Staats & Staats, 1958), mere exposure effects (Zajonc, 1968), balance and congruity theories (e.g., Heider, 1958; Osgood & Tannenbaum, 1955), cognitive dissonance (Festinger, 1957, 1964), and social judgment theory (Sherif & Hovland, 1961), among others (see Petty & Wegener, 1998).

By the late 1960s and early 1970s, however, the attitudes domain faced criticism on three fronts. First, critics argued that attitudes often were unrelated or only weakly related to behavior (Wicker, 1969). Second, researchers mused that changes of attitudes in the laboratory often failed to persist outside the lab (e.g., Cook & Flay, 1978). Finally, theorists noted that few generalizable principles of attitude change had been identified in the 30 years between the late 1940s and the late 1970s (e.g., Sherif, 1977). In a sense, these theorists were noting a replication crisis of sorts in which certain persuasion effects, such as credible sources resulting in greater persuasion, occurred in some studies but not in others or occurred using some message topics but not others. As discussed in greater detail in later sections of the chapter, however, these challenges spawned theoretical developments that have organized and guided research on attitudes ever since the criticisms were leveled.

A third era in attitudes research built off of the challenges of the 1970s to develop theories that brought order to the seeming chaos of findings. Starting in the mid-1970s and fully under way by the early 1980s, there was substantial development of dual- and multiprocess theories that specified when particular persuasion effects would be most likely to occur, when attitude changes would be most likely to last over time, and when attitudes would be most likely to guide future behavior. In particular, the elaboration likelihood model (ELM; Petty & Cacioppo, 1986) and heuristic-systematic model (HSM; Chaiken, Liberman, & Eagly, 1989) suggested that the relative impact of "central" features of an attitude object versus more "peripheral" aspects of a persuasive appeal depends on the extent to which one is thinking deeply about the attitude object. That is, using ELM terminology, a key moderator of persuasion effects is the extent to which one is elaborating on attitude-relevant information—where elaboration involves scrutinizing the merits of the attitude object by going beyond the presented information to compare it with relevant knowledge stored in memory. (See Petty & Cacioppo, 1986, and Wegener & Petty, 2013, for additional discussion of links with the concept of elaboration in cognitive psychology.) Although originally more aligned with the message learning tradition, the HSM concept of systematic processing has been treated in ways largely similar to how elaboration has been treated (Chaiken et al., 1989). Research inspired by both the ELM and HSM has shown, for example, that qualities of the source of a message (such as the source's credibility) can influence postmessage attitudes more when elaboration or systematic processing is relatively low rather than high (e.g., Chaiken, 1980; Petty, Cacioppo, & Goldman, 1981). In contrast, manipulations of the quality of the arguments in a persuasive message generally influence attitudes more when motivation and ability to elaborate are high rather than low (e.g., Petty et al., 1981). Elaboration also has been tied to persistence of attitudes over time and to the likelihood of attitude-consistent action (Petty, Haugtvedt, & Smith, 1995). In addition to the ELM and HSM, Fazio's (1986, 1990) MODE (*M*otivation and *O*pportunity as *De*terminants) model of attitude–behavior processes systematically addressed moderators of when attitudes are most likely to guide behavior.

The 1980s and 1990s represented a fundamental shift in the zeitgeist, from addressing attitudes as somewhat static objects of measurement and unidirectional influence to a focus on larger theories of the underlying, structural principles of attitudes. That is, the concept of elaboration suggested an enriched cognitive structure underlying the attitude, and the MODE model focused on accessibility of attitudes in memory. (See the discussion of properties of attitudes related to attitude strength in the section titled "Attitude–Behavior Consistency.") In addition, although raised in previous eras, the study of attitude ambivalence (having a mixture of both positive and negative reactions to an attitude object) was rejuvenated in the 1990s (e.g., Cacioppo & Berntson, 1994; Priester & Petty, 1996). Such structural emphases also played key roles in development of new indirect attitude measures that further contributed to studies of hypothesized structure-based influences of attitudes. (See the discussion of indirect attitude measures based on reaction times in the subsection titled "Measurement: Priming and Response Time Measures.") Such emphases carried through the early 2000s and continue to receive a good deal of research attention (e.g., see Fabrigar et al., 2005; Petty, Fazio, & Briñol, 2009).

Although it is likely too soon to identify the next era in attitudes research, a contending theme that we discuss later in the chapter, in the subsection titled "Metacognition and Attitude Properties," is that of thoughts about one's thoughts and thought processes (i.e., metacognition; Petty, Briñol, Tormala, & Wegener, 2007).

Functions of Attitudes

Attitudes facilitate adaptations to the environment (Eagly & Chaiken, 1998). They prepare a person for action or for interaction with his or her decision options. Seminal work by Daniel Katz (1960) outlined four common attitudinal functions that remain generally accepted in the field today. The *utilitarian function* helps people to gain rewards and avoid punishments. That is, attitudes help people to reach their personal goals or desires (Carpenter, 2012). For example, liking the same band as a potential romantic partner might help a person maintain the relationship and avoid rejection. The *value-expressive function* helps people express central, unique values. That is, attitudes often are maintained in an effort to remain consistent with core values or beliefs. Rejecting a diamond proposal ring, not on the basis of any negative feelings toward the courter, but because one holds negative attitudes toward unethical mining practices, would be an example of an attitude being used in a value-expressive fashion. Attitudes serving an *ego-defensive* function aid the person in feeling good about him- or herself (i.e., protecting one's self-esteem). That is, attitudes may be used to rationalize failure or void threats to self-esteem, such as from insults. The literature on the self is rife with research on methods people use to preserve their self-esteem, but in Katz's initial conception (Katz, Sarnoff, & McClintock, 1956), these attitudes were thought often to involve a sense of superiority over other individuals. Attitudes also can serve a *knowledge function*. That is, attitudes allow people to make sense of the world by helping them to organize information and interpret novel information. In other words, attitudes can help to give people a sense of ability to predict and control their environment (cf. Heider, 1958).

These functions may be considered as representing major motivational underpinnings for the use of attitudes. Some objects may be generally associated with a given attitude function (Shavitt, 1990). However, the same attitude may serve a utilitarian function for one person and a value-expressive function for another (for variations based on individual differences, see DeBono & Packer, 1988; Snyder & DeBono, 1985) or even different functions for the same person depending on the circumstances (e.g., Fazio, Lenn, & Effrein, 1984). Although some combination of these four functions is widely accepted by most researchers in the literature, they should not be considered to be exhaustive. For instance, research suggests that attitudes additionally may hold a social-adjustive function (M. B. Smith, Bruner, & White, 1956), in which they are used to assist people in fitting into groups and functioning in a social world.

On a process level, attitudes have been shown to create attitude-consistent biases in perception (e.g., Hastorf & Cantril, 1954), information processing (e.g., Lord, Ross, & Lepper, 1979; Munro & Ditto, 1997), classification of objects and related judgments (e.g., E. R. Smith, Fazio, & Cejka, 1996), as well as exposure to (and memory for) attitude-relevant information (e.g., S. M. Smith, Fabrigar, Powell, & Estrada, 2007; but see Eagly, Kulesa, Brannon, Shaw, & Hutson-Comeaux, 2000, for mechanisms that increase memory of attitude-inconsistent information).

Perhaps the most important function of attitudes, however, may be the role they play in human behavior. William James famously wrote, "My thinking is first and last and always for the sake of my doing" (James, 1890/1950, p. 333). If thinking is for doing, then attitudes are for action. Marketing firms invest a tremendous amount of resources in order to understand the preferences of their consumers: whether they prefer red or blue packaging, whether they will pay more for a deluxe service, whether they like the taste of some brands over others. This is not intended to be an exercise in futility but as a method of forecasting behaviors, such as how many people will purchase a new smartphone (based on their attitudes toward the phone, the company, the advertising, etc.). From ads to political polls, attempts at adjusting, detecting, and measuring attitudes are ubiquitous. But the attitude itself is often not the end goal of the marketer, the pollster, or the researcher: The action that will follow is.

In sum, one of the primary reasons people care about attitudes is that they inform action. What we like or dislike determines what we do. However, as in any psychological domain, the full account of human behavior is much more complicated than a simple link between attitudes and behavior. Understanding *when*, *why*, and *how* attitudes influence behavior is essential to allowing predictions from any given attitude measure to any particular behavior.

Attitude–Behavior Consistency

Against the backdrop of behavior prediction being a primary reason to study attitudes, studies that produced little to no relation between attitudes and behavior were more than concerning. Yet a meta-analysis of 31 studies by Wicker (1969) found dismal relations between attitude measures and behaviors, leading Wicker (p. 75) to state that "[t]he present review provides little evidence to support the postulated existence of stable, underlying attitudes within the individual which influence both his verbal expressions and his actions." Following closely on the heels of this work, a paper by Abelson (1972) titled "Are Attitudes Necessary?" questioned the fundamental nature of attitudes as a psychological construct.

Such challenges to the utility of attitudes in predicting behavior spawned three classes of methodological and theoretical developments. First, researchers examined the extent to which measures of attitudes and behaviors seemed to be focused on the same action, target, context, and time. A second response to these concerns specified the mediational relations (underlying mechanisms) between attitudes and behaviors and also introduced alternative predictors that might influence similar mediators and therefore impact the attitude–behavior relation. Finally, research began to identify properties of attitudes that help determine the extent to which the attitudes are capable of guiding behavior. This moderational approach developed into what many researchers would consider the *attitude strength* approach to attitude–behavior consistency.

Correspondence of Attitude and Behavior Measures

Ajzen and Fishbein (1977) suggested that attitude–behavior relations would be stronger to the extent that measures of the attitude and behavior both take into account the following: (1) the action performed (e.g., donating blood), (2) the target at which the action is directed (e.g., to the Red Cross), (3) the context in which the action is performed (e.g., a local blood donation center), and (4) the time (e.g., Thursday at 5 pm). Attitudes should not predict behaviors when the attitudes address a target other than the target of the behavior (as Ajzen and Fishbein suggested was the case in many of the studies reviewed by Wicker). Ajzen and Fishbein suggested that each level of correspondence or "compatibility" in the elements addressed by each measure should increase the observed relations between attitudes and behaviors. This means that specific attitudes should predict specific behaviors, whereas general attitudes should predict general (aggregated) behaviors. Ajzen and Fishbein classified attitude–behavior studies according to the number of dimensions (action, target, context, and time) along which the measures of attitude and behavior were compatible. For example, if attitude and behavior measures both assessed blood donation (action) to the Red Cross (target) at a local blood donation center (context), then these measures would be said to be compatible on three dimensions. In Ajzen and Fishbein's review, every study with attitude and behavior measures compatible on three or more dimensions showed significant attitude–behavior relations. In contrast, studies using measures with less compatibility across dimensions generally showed nonsignificant and weak relations.

A number of studies directly demonstrated the influences of compatibility of attitude and behavior measures. For example, Davidson and Jaccard (1979) measured attitudes in a number of ways that differed in the extent to which they were compatible with the behavioral measure of women's use of birth control pills over the 2-year period of the study. When initial attitude measures simply asked for attitudes toward birth control, the correlation between attitudes and behavior was quite weak ($r = .08$). The relation was stronger ($r = .3$) when attitudes were measured toward birth control pills and stronger yet when attitude measures examined attitudes toward using birth control pills in general ($r = .5$) or during the next 2 years ($r = .6$). Thus, consistent with the Ajzen and Fishbein (1977) analysis, making the attitude object more specifically match the behavior of interest increased the relation between the attitudes and behaviors. Such effects show that specific attitudes can relate to specific behaviors, but many researchers have been interested in attitudes precisely because of their general nature. Recall, however, that the Ajzen and Fishbein analysis did not suggest that more general attitudes were incapable of predicting behavior. Rather, they suggested that general attitudes should do a better job of predicting more general indices of behavioral tendencies across varied settings (rather than predicting isolated behaviors in specific settings).

Weigel and Newman (1976) examined people's general attitudes toward environmental preservation and provided residents with a number of behavioral opportunities (such as signing pro-environment petitions, engaging in roadside cleanup efforts, and recycling). Attitude–behavior relations were relatively weak when examining each specific behavior (median $r = .3$), but the relation was much stronger when examining an index created across all the behavioral opportunities ($r = .6$). Thus, general attitudes predicted a general index of behavior. Overall, the work inspired by Fishbein and Ajzen's (1977) comments about attitude and behavior measurement made a strong case that attitudes can be strongly related to behavioral outcomes, given compatible measures of each construct.

Mediational Mechanisms

Another approach to low attitude–behavior relations focused on mechanisms by which attitudes predict behavior, noting that attitudes often might work through other constructs and that predictors other than attitudes per se also contribute. The theory of reasoned action (Fishbein & Ajzen, 1975) and theory of planned behavior (Ajzen, 1985) have been hugely influential in this area.

In both models, people's attitudes toward a behavior are one determinant of behavioral intentions, and behavioral intentions serve as a proximal predictor of behavior. Thus, an attitude might not relate closely to behavior if the attitude fails to create an intention to act in a particular way or if the ultimate action does not reflect the intention. An important aspect of this approach is that attitudes are not the only influences on intentions. That is, subjective norms (i.e., what other important individuals believe one should do) and perceptions of control over the behavior in question (in the theory of planned behavior) also influence intentions. Thus, a second reason attitudes sometimes fail to influence behavior is that subjective norms or perceived behavioral control are driving the intentions. The behavioral intentions of some individuals are more attitudinally driven, whereas the intentions of other individuals are more normatively driven (Trafimow & Finlay, 1996). Similarly, the activation of different self-views can change the extent to which behavioral intentions are based on attitudes versus subjective norms (Ybarra & Trafimow, 1998).

Moderators of Attitude–Behavior Consistency

A third approach to (lack of) attitude–behavior consistency was to attempt to identify properties of attitudes that were most likely to guide behavior. This *moderation* approach arguably has had the most influence on the attitudes literature, as many of the qualities of attitudes also have been studied in relation to alternative outcomes, such as resistance to persuasion or influences on perception and judgment. Unlike the measurement compatibility and mediation approaches, most research on attitude properties has examined attitudes toward targets of behavior rather than toward the behaviors themselves.

Attitude Strength. Research identified many properties of attitudes that increase attitude–behavior consistency. For example, attitudes predict behavior better when the attitude is based on direct experience with the attitude object (e.g., Regan & Fazio, 1977). Attitudes also better predict behavior or behavioral intentions when the attitude is based on relatively high levels of elaboration (e.g., Cacioppo, Petty, Kao, & Rodriguez, 1986), when the attitude is important to the person (Boninger, Krosnick, Berent, & Fabrigar, 1995), when the attitude is held with certainty (Tormala & Rucker, 2007), when the attitude is unambivalent (Thompson, Zanna, & Griffin, 1995), and when knowledge about the attitude object is high (Wood, Rhodes, & Biek, 1995). These predictors of attitude strength can be related, but often not closely (e.g., Krosnick, Boninger, Chuang, Berent, & Carnot, 1993). These moderators can be objective or subjective, even for the same attitude property (cf. Bassili, 1996). For example, one may be objectively ambivalent if one acknowledges both positive and negative aspects of the object and may be subjectively ambivalent if one feels mixed or torn about the attitude object. Subjective perceptions of strength-related properties can stem from the objective qualities of the object (e.g., Barden & Petty, 2008; Priester & Petty, 1996) but also can be influenced separately (e.g., by comparisons with other people's evaluations, Priester & Petty, 2001).

In general, these properties of attitudes have been considered either as antecedents of attitude strength (when strength is defined with impact on behavior as one indicator; e.g., Krosnick & Petty, 1995) or as definitive of strength itself (in the case of attitude accessibility, where the strength of the association in memory between the attitude object and the evaluation—the accessibility of the attitude—is treated as the same concept as attitude strength; e.g., Fazio, 1995).

MODE Model. Fazio (1990) treated the theory of reasoned action as mostly addressing behaviors that are deliberately considered. In contrast, many behaviors are rather spontaneous. The MODE model delineated relatively spontaneous and deliberative determinants of behavior and, in so doing, has become a highly influential model. As its name implies, the MODE model suggests that motivation and opportunity to deliberate determine the mechanisms that influence behavior. In various MODE-inspired studies, the impact of overall attitudes is greater when motivation or opportunity to deliberate is low, especially when attitudes are highly accessible in memory (e.g., Schuette & Fazio, 1995). When motivation or opportunity to deliberate is lacking, one's attitudes may color one's perceptions of a given target in an attitude-congruent manner (e.g., viewing a liked politician especially positively during a debate; Fazio & Williams, 1986; Houston & Fazio, 1989). This attitudinal tinting of one's perceptions may in turn lead to attitude-consistent behavior, if one is not motivated and able to avoid the impact of one's attitudes. Sometimes such motivation might stem from having alternative action-related information that is viewed as more diagnostic than the attitude (e.g., Sanbonmatsu & Fazio, 1990), but in other cases, the motivation might stem from a wish not be influenced by the attitude (e.g., wanting not to appear prejudiced; Fazio et al., 1995).

BEYOND BEHAVIOR: OTHER IMPACTS OF ATTITUDES

For many researchers, attitude–behavior consistency is only one indicator of attitude strength. Other indicators include an attitude's ability to guide information processing, to resist persuasion, and to remain stable over time (e.g., Krosnick & Petty, 1995). This conceptualization treats attitude strength as analogous to the concept of strength in a muscle, where the muscle's strength is characterized by its ability to last over time (slow to atrophy), to exert influence (ability to push or pull), and to resist movement (to withstand outside pressure). Consistent with the notion of multiple indicators of attitude strength, many of the antecedent properties that have been studied as moderating attitude–behavior relations have been found to influence these other strength-related outcomes. For example, influences of attitudes on information processing and selective exposure to information have been moderated by attitude accessibility (e.g., Houston & Fazio, 1989) and attitude certainty (e.g., Sawicki et al., 2011). Resistance to change has been affected by the amount of elaboration on which the attitude is based (e.g., Haugtvedt & Petty, 1992; Haugtvedt & Wegener, 1994), the level of attitude accessibility (e.g., Bassili, 1996), the amount of knowledge associated with the attitude (e.g., Wood, 1982), and the amount of attitude ambivalence (e.g., Armitage & Conner, 2000). Also, persistence of attitudes over time has been associated with direct experience (e.g., Doll & Ajzen, 1992), elaboration (e.g., Haugtvedt & Petty, 1992), certainty (e.g., Pelham, 1991), and importance (e.g., Krosnick, 1988).

Attitudes are consequential in determining not only what we do but also how easily we do it (e.g., Fazio, Blascovich, & Driscoll, 1992), how we view a given object (Lord et al., 1979), and even whether we are likely to notice that object in the first place (Roskos-Ewoldsen & Fazio, 1992). Given the consequential nature of attitudes, it should come as no surprise that attitude change, or *persuasion*, also has a long research history.

PERSUASION: ATTITUDE CHANGE

Brief Overview of Conceptual Developments

Message Learning

An understanding of persuasion is important for "basic" researchers and "applied" practitioners alike. Following World War II, Carl Hovland assembled an active team of researchers to study persuasion systematically. The team's message learning approach (Hovland et al., 1953) treated persuasion like skill acquisition with attention, comprehension, yielding, and retention phases. That is, the effectiveness of a persuasive message was assumed to depend on sufficient attention to the message, understanding of the message, taking on the conclusion of the message as one's own view of the attitude object, and retention of the information in the message and its conclusion. After repeated attempts failed to show relations between memory for message content and postmessage attitudes (e.g., Insko, 1964; Miller & Campbell, 1959), interest waned in message learning per se. Yet theoretical developments built on the message learning approach to start treating message recipients as more active responders to persuasive messages. For example, McGuire (1968) treated yielding as separable and often more active than reception of a message (i.e., attention, comprehension, and retention), and he identified variables that might increase yielding but decrease reception (or vice versa). Contemporary research suggested that self-generated arguments were more effective than passively received arguments (Janis & King, 1954), that active persuasion lasted longer than passive persuasion (Watts, 1967), and that message recipients' cognitive responses to a persuasive message served as the key determinant of attitude change (e.g., Greenwald, 1968).

Cognitive Dissonance Theory

As research transitioned away from message learning, research accelerated on cognitive consistency pressures that could make attitude-inconsistent behavior change the associated attitudes. In particular, the theory of cognitive dissonance proposed that two inconsistent cognitions, such as knowing that smoking is deadly but also knowing that one continues to smoke, lead to a state of unpleasant arousal that one is motivated to reduce (Festinger, 1957, 1964). One may reduce this unpleasant arousal by trivializing one of the cognitions (e.g., "The harmful effects of smoking aren't a big deal to me"), adding a new consonant cognition that reconciles the formerly dissonant cognitions (e.g., "I eat healthily, so that will 'make up' for my smoking habit"), or changing a relevant attitude or behavior to make the two cognitions consistent (e.g., becoming more in favor of smoking after the belief-inconsistent action of smoking created discomfort). The mode of dissonance reduction can depend on ease of resolution through that means (Abelson, 1959) or on which mode presents itself first (e.g., Simon, Greenberg, & Brehm, 1995).

Dissonance theory helped predict changes in attitudes across a wide variety of behavior settings. For example, dissonance theory explained why severe initiation can lead to greater liking for a group—because a favorable attitude justifies the effort (Aronson & Mills, 1959). Receiving insufficient justification for engaging in counterattitudinal behavior (e.g., receiving a small amount of money to lie to another person about a task that person will undertake) can change attitudes to become more consistent with the action (Festinger & Carlsmith, 1959). Similar insufficient justification rationale predicted that children avoiding a desirable toy because of mild rather than severe threats would come to view the ignored toy less favorably (e.g., Aronson & Carlsmith, 1962). Dissonance also explained postchoice spreading of alternatives in which the chosen option is viewed more positively than it was pre-choice and the rejected option is viewed less positively than it was pre-choice (Brehm, 1956; but see Chen & Risen, 2010).

Dissonance versus Self-Perception. The dissonance perspective did meet with skepticism. Consistent with more cognitive perspectives in the 1970s, self-perception theory proposed that people sometimes infer their attitudes from their own behavior, much like people infer the attitudes of others from others' behaviors (e.g., Bem, 1972). Because contexts for the behaviors would be taken into account in such inferences, the two theories seemed to make the same predictions (Greenwald, 1975). The theories did differ, however, on the hypothesized processes leading to the effects, especially the unpleasant tension that was supposed to motivate dissonance-based attitude change. Although early efforts to measure unpleasant arousal were inconclusive, Zanna and Cooper (1974) used a misattribution paradigm to show that unpleasant arousal was involved in dissonance-based attitude change produced by freely writing a counterattitudinal essay. When uncomfortable tension could be attributed to a pill, postessay attitude change was reduced; but when the pill was thought to be relaxing, postessay attitude change actually increased. Later attempts to measure dissonance-based discomfort also produced support for heightened autonomic arousal (e.g., Elkin & Leippe, 1986), though the observed attitude change was shown to be motivated by negative affect rather than arousal per se (e.g., Losch & Cacioppo, 1990; see also Elliot & Devine, 1994). Such studies provided strong evidence that at least some dissonance-inspired effects were motivated by the unpleasant state that dissonance theory predicted. However, the presence of aversive arousal in some settings did not mean that self-perception never occurs. For example, Fazio, Zanna, and Cooper (1977) suggested that self-perception and dissonance might have different domains of operation, with self-perception capable of influencing attitudes when the behavior is not viewed as objectionable (thereby reducing any discomfort).

Aversive Consequences. Many moderators of dissonance effects seemed to go beyond mere inconsistency between cognitive elements. For example, behavior should have been dissonant from beliefs when telling a next participant that a boring task was interesting (Festinger & Carlsmith, 1959). Yet the attitude change was observed only when the next participant believed the statement (e.g., Calder, Ross, & Insko, 1973). Similarly, freely choosing to write a counterattitudinal essay should be dissonant (Zanna & Cooper, 1974), but attitude change only occurred when a decision-making body was to use the essay as support for an unwanted decision (Collins & Hoyt, 1972). These types of outcomes prompted Cooper and Fazio (1984) to suggest that dissonance starts with an aversive consequence of an action, followed by an assessment of whether one is personally responsible for the aversive consequence and whether the consequence was foreseeable. If each of these is true, then dissonance will be experienced.

Self in Dissonance. Early in the development of dissonance theory, Aronson (1969) suggested that dissonance really stemmed from an inconsistency between an undesirable behavior and a positive self-concept. For example, lying to another person about a boring task would be dissonance-provoking only if the actor views him- or herself as honest. This self-consistency view predicted that dissonance might be more pronounced for people who hold more positive self-views (e.g., those with high self-esteem; Cooper & Duncan, 1971). Interestingly, self-affirmation theory (Steele, 1988) made the opposite prediction. Similar to the self-consistency view, self-affirmation theory proposed that dissonance occurs because a negative behavior threatens a person's positive self-views. But in self-affirmation theory, the key is a general sense of integrity and self-worth rather than specific beliefs about traits that diverge from the implications of a behavior. Thus, in self-affirmation theory, any means of regaining one's sense of self-worth is sufficient to reduce dissonance motives, and that is true regardless of whether the regaining of self-worth directly addresses the inconsistency that initially created the motive. For example, Steele and Liu (1983) reduced dissonance effects in the counterattitudinal essay paradigm when research participants were reminded of an unrelated but valued aspect of themselves before reporting their postessay attitudes. In self-affirmation theory, then, self-esteem serves as a resource that can buffer against self-threats. Thus, high self-esteem would reduce (rather than enhance) the effects of typical dissonance inductions (e.g., Steele, Spencer, & Lynch, 1993).

Dual-Process and Multiprocess Theories

Elaboration Likelihood Model and Heuristic-Systematic Model

As noted earlier, crises in the 1970s related not only to attitude–behavior relations but also to consistency in persuasion effects and persistence of attitude effects over time. Some researchers viewed the diverse, sometimes contradictory effects as evidence of a dying field, but others regarded the confusion as an opportunity to develop new theoretical models that would integrate the various processes and identify conditions under which different effects should occur. The elaboration likelihood model (Petty & Cacioppo, 1979, 1986) and heuristic-systematic model (Chaiken, 1980; Chaiken et al., 1989) each proposed that the persuasion process at work would depend on the extent to which a message recipient was motivated and able to think carefully about the attitude object. These theories explained when and how various persuasion effects are most likely to occur (e.g., when credible sources lead to more, less, or equal persuasion compared with noncredible sources). The ELM also made predictions linking the amount of elaboration to the resulting strength of the attitude. The HSM was explicitly a dual-process theory, juxtaposing heuristic processing with systematic processing. In contrast, the ELM was developed as a multiprocess model, with various types of relatively thoughtful or nonthoughtful processes occurring across an elaboration continuum. In the remainder of this chapter, we use the ELM language, but many of the primary persuasion effects in question when the models were developed would be predicted equally well by the two models. (For comparison of the two models, see Chen & Chaiken, 1999; Petty & Wegener, 1999).

Elaboration Continuum

The ELM conceptualizes persuasion processes as lying along an elaboration continuum. The ELM built on the differences between relatively passive message recipients in the message learning tradition and more active message recipients in later treatments, such as the cognitive response approach. In particular, the ELM noted that motivation and ability to elaborate vary across people and situations. As motivation and ability increase, the impact of the perceived central merits of the attitude object also should increase, but if motivation or ability is lacking, more peripheral aspects of a persuasive communication can affect attitudes. In many ELM-inspired persuasion studies, the central merits of the attitude object are manipulated through the quality of arguments provided in the persuasive message. One or more peripheral aspects of the communication, such as who presents the information or the background music accompanying a message, might also be manipulated. Thus, for example, high levels of motivation to elaborate have increased the impact of argument quality on persuasion, whereas lower levels of motivation have led source characteristics, such as source expertise or source attractiveness to have greater impact (e.g., Petty et al., 1981). Individual differences in motivation to elaborate also have been identified (i.e., need for cognition; Cacioppo & Petty, 1982; Cacioppo, Petty, Feinstein, & Jarvis, 1996). When sufficient motivation exists, differences in ability to elaborate also move people higher or lower on the elaboration continuum. That is, higher levels of ability lead to larger argument quality effects (e.g., Petty, Wells, & Brock, 1976) but smaller effects of peripheral cues (e.g., Kiesler & Mathog, 1968).

Early flowchart depictions of the model labeled the endpoints of the elaboration continuum as two "routes" to persuasion (i.e., the central route and peripheral route). It is important to realize, though, that these labels were simply endpoints of an underlying elaboration continuum. Across that continuum, varying amounts of impact of central merits and peripheral cues would be expected, so that many points along the continuum would produce some impact of both arguments and cues (see also Petty & Wegener, 1999).

Multiple Roles for Persuasion Variables

Although certain central merits might be object-specific, many persuasion variables, such as endorser attractiveness, sometimes can serve as a peripheral cue and sometimes can serve as a central merit (i.e., as a primary feature of a high-quality version of the attitude object). When the variable is easily processed and perceived as relatively unrelated to the central merits of the attitude object, it could serve as a cue and have greater influence on postmessage attitudes when motivation or ability to process is lacking (i.e., at the low end of the elaboration continuum). However, when the variable is perceived as a central merit of the attitude object and it takes some processing to assess the level of the variable, it generally should have greater influence on postmessage attitudes when motivation and ability to process are relatively high (i.e., at the high end of the elaboration continuum). Serving as cues or arguments are not the only types of effects that persuasion variables might have, however. Thus, the ELM specified multiple roles for persuasion variables across the elaboration continuum. Each role is connected to a particular portion of the continuum. As noted earlier, argument effects should be more likely toward the high end of the continuum, and cue effects should be most likely at the low end of the continuum. When motivation to elaborate is unclear (e.g., the person is unsure whether the issue is worth the processing effort), however, the same variables might influence the extent to which people elaborate. For example, people might be more motivated to process a message from an expert rather than from a nonexpert source (Heesacker, Petty, & Cacioppo, 1983).

The cue, argument, and amount of processing roles generally are identified when arguments in a message are clearly strong (compelling) or weak (specious). In some settings, though, available arguments are somewhat ambiguous or vague. In such settings, high levels of elaboration can lead to biases in processing by the same persuasion variables that played the other roles under the specific conditions outlined earlier. For example, research by Chaiken and Maheswaran (1994) replicated the cue effects of source credibility from Petty et al. (1981) when clearly strong or weak arguments were used. Yet when arguments were less clear (a mixture of weak and strong arguments), source expertise biased elaboration under high-motivation conditions.

Multiple roles have been examined for a number of variables. Beyond source expertise, multiple roles have been examined for source attractiveness (e.g., Shavitt, Swan, Lowrey, & Wänke, 1994) and source efficacy (e.g., Clark, Evans, & Wegener, 2011; Clark & Wegener, 2009). Perhaps the most thoroughly examined variable for multiple roles has been message recipient mood. Mood has been examined as a persuasion cue in low-elaboration conditions (Petty, Schumann, Richman, & Strathman, 1993), as influencing amount of processing in more moderate elaboration conditions (Wegener, Petty, & Smith, 1995), and as biasing processing or serving as an argument in high-elaboration conditions (Martin, Abend, Sedikides, & Green, 1997; Petty et al., 1993). In such cases, the ELM does not specify whether, for example, happy mood should increase or decrease processing (or should create positive versus negative biases in processing). In fact, different theories have been developed to inform the processing effects of mood (e.g., compare Wegener et al., 1995, and Ziegler, 2014) and moderators of the direction of biases in processing (e.g., Wegener, Petty, & Klein, 1994). In each case, however, the ELM set parameters on the levels of elaboration at which each type of effect should be most likely to occur.

The Elaboration Continuum as an Organizing Framework

On a meta-theoretical level, the elaboration continuum can help organize diverse persuasion theories and mechanisms in terms of the amount of elaboration that they require (Petty & Wegener, 1998). This organization helps to clarify how different models or processes may not be contradictory but rather may occur under different circumstances and, in a sense, occupy different domains. Next we classify various persuasion theories and processes based on where they likely fall on an elaboration continuum. For more in-depth discussions, see Petty and Wegener (1998) and Wegener and Carlston (2005).

Persuasion Under Relatively Lower Levels of Elaboration

Conditioning. A variety of types of conditioning might influence attitudes without much need for elaboration. Indeed, early research applied classical conditioning principles to attitudes, where an initially neutral stimulus, such as an uncommon country name, served as the conditioned stimulus, and clearly positive or negative words served as the unconditioned stimuli in a higher-order conditioning paradigm (e.g., Staats & Staats, 1958). Classical conditioning involves signal learning in which the link between the unconditioned stimulus and the conditioned stimuli may be known. Evaluative conditioning describes situations in which presentation of a valenced (positive or negative) stimulus leads to that valence "bleeding" onto an attitude object—a form of persuasion that does not require contingency awareness (i.e., awareness of the pairing; see, e.g., Baeyens, Eelen, & Van den Bergh, 1990; Dijksterhuis, 2004). Just as one's behaviors may be shaped through reinforcement and punishment, so too may one's attitudes. In one classic demonstration of operant conditioning of attitudes, Hildum and Brown (1956) found that a telephone interviewer could shape interviewees' attitudes through selective verbal reinforcement. Operant conditioning has limitations, such as the requirement that the recipient have some behavior or response to be reinforced or punished. At any rate, various forms of conditioning can influence attitudes with little need for active elaboration of attitude-relevant information.

Mere Exposure. People tend to develop more favorable attitudes toward an object as they become more familiar with it—even when "merely exposed" to the object. The effect has been shown with many different kinds of stimuli, such as research-confederate "students" attending classes but not interacting with other students (Moreland & Beach, 1992) and briefly or even subliminally presented Chinese ideographs or abstract patterns (Bornstein & D'Agostino, 1992; Monahan, Murphy, & Zajonc, 2000). Various theories account for the mere exposure effect. For example, perceptual fluency of the repeated stimuli may be experienced positively (Winkielman & Cacioppo, 2001). Alternatively, the ease of perceiving the object may be misattributed to liking of the object (Bornstein & D'Agostino, 1992). Most pertinent to the current discussion, mere exposure may be greatest when one is not elaborating thoroughly (see Kruglanski, Freund, & Bar-Tal, 1996).

Use of Heuristics. One way to account for many cue effects is to reference rules of thumb known as heuristics (e.g., "I should trust a likable source"). Application of heuristics can influence persuasion, especially under low-elaboration conditions (e.g., Chaiken, 1980).

Balance and Congruity. Balance theory (Heider, 1958) and congruity theory (Osgood & Tannenbaum, 1955) posit that people like to feel consistent (or "balanced") by agreeing with those whom they like and disagreeing with those whom they dislike. Thus, when people learn of a liked person's opinion, they would rather agree than disagree with the person (Osgood & Tannenbaum, 1955), and when people learn that they agree rather than disagree with a stranger, they come to like the stranger (Byrne, 1971). Although balance-based persuasion involves some knowledge of another person's attitude (as well as one's attitude toward that other person), it does not require consideration of central merits of the attitude object in question. Therefore, it seems that balance or congruity effects could occur at relatively low levels of elaboration.

Persuasion Under Relatively Higher Levels of Elaboration

Message Learning and Revisions. As described earlier, the learning approach taken by the Yale Group generally relied on notions of rehearsal rather than elaboration. Even so, paying attention to a message and working to comprehend it might be considered as taking more than minimal effort in thinking. Also, especially in some revisions of the message learning approach, processes more like elaboration were involved. For example, in McGuire's (1968) reception and yielding approach, yielding often consisted of cognitive responses that were favorable rather than unfavorable toward the message or its conclusion. Recipient variables such as self-esteem and intelligence might be positively associated with reception (e.g., intelligent people might be more capable of comprehending a message) but negatively associated with yielding (e.g., intelligent people might be more likely to actively oppose a persuasion attempt). As a result, unless the qualities of the message place an emphasis on only reception or only yielding, moderate levels of the recipient variables produce maximal persuasion (see Rhodes & Wood, 1992).

Cognitive Responses to Messages. The original cognitive response approach emphasized the favorability of individuals' thoughts in response to a message, independent of whether those thoughts were "correct" or even central to the message (Greenwald, 1968). However, when restricted to thoughts about the attitude object and its central merits, cognitive responses and attitudes tend to correlate most strongly for high-involvement issues, when one presumably is engaging in the most elaboration (Petty & Cacioppo, 1979).

Mere Thought. Merely thinking about a topic (in the absence of a message) can lead to attitude polarization (e.g., Tesser, 1976). This mere thought effect results from multiple mechanisms. The more attitude-congruent thoughts there are, the more polarization exists. Mere thought also increases confidence in one's thoughts, which in turn leads to attitude polarization (Clarkson, Tormala, & Leone, 2011). This finding suggests that mere thought sometimes may backfire: When people have an excessive amount of time during which to think, they eventually may have difficulty coming up with new thoughts. This difficulty can reduce confidence in one's thoughts, thereby blunting attitude polarization (Clarkson et al., 2011).

Expectancy/ Value Approaches. Fishbein and Ajzen (1975) conceptualized attitudes as composed of sums of expectancy × value components, where expectancies (the likelihood that the attitude object possesses some characteristic) were multiplied by values (the desirability of the characteristic). Attitude change, in turn, may operate via changes in expectancies or changes in values of characteristics associated with the attitude object. For example, positive mood sometimes may lead to persuasion by increasing the perceived likelihood of a favorable outcome attributed to a policy, and these processes may be especially likely under high-elaboration conditions (e.g., Wegener et al., 1994).

Cognitive Dissonance. There is some ongoing debate concerning whether cognitive dissonance involves elaboration. At least some dissonance paradigms and some means of dissonance reduction (such as generation of consonant cognitions or thinking biased by the motive to reduce the discomfort of dissonance) seem likely to involve some level of elaboration. The facts that some dissonance effects last over time (e.g., Freedman, 1965) and appear primarily on "propositional" rather than merely associative outcome measures (e.g., Gawronski & Strack, 2004) are consistent with that idea.

RECENT/ EMERGING TRENDS

Measurement: Priming and Response Time Measures

In recent years, the vast majority of research using indirect measures has made use of some type of "implicit" measure. These measures rely on brief presentations of attitude objects and (quick) responses to (often unrelated) evaluative questions. The measures are indirect in that they do not ask participants to evaluate the attitude object. The responses also are assumed to be less controllable than with direct measures, and at least sometimes the measures have been assumed capable of tapping into evaluative associations of which respondents are unaware. The basis for these assumptions might be debatable (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Gawronski, Hofmann, & Wilbur, 2006) and might apply differentially to the different measures. However, no one can question the popularity of the measures. They have facilitated a wide variety of research questions that would have been unlikely using direct measures alone. Such questions include the role of propositional versus associative processes in attitude change (Gawronski & Bodenhausen, 2006; see also Gawronski & Payne, 2011; Petty et al., 2009). The first two measures, the evaluative priming measure and the Implicit Association Test (IAT), are reaction time measures. At the core of these measures is the idea that people have mental associations between an object and their evaluation of it. Generally, people should be faster to categorize positive objects as positive if those people are first "primed" with positivity rather than negativity, but response competition should slow responses when primed valence mismatches with the valence of the object to be judged. This general idea is implemented in different ways in each measure. The third measure, the affect misattribution procedure (AMP), is not based on reaction time, but it uses judgments of unrelated stimuli that follow brief exposures of the attitude object.

Evaluative Priming and the IAT

In the evaluative priming procedure, an attitude object is presented prior to an unrelated target adjective that is categorized as either positive or negative as quickly and accurately as possible (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Speeded responding toward the target word is taken as an indication that the spontaneous evaluation of the prime matched the valence of the target. Moreover, because the prime is presented very briefly and the time between appearance of the prime and appearance of the target word is too fast for effortful thought (e.g., 350 ms), influences of the prime on the speed of target responses often are taken as indicating automatic activation of the attitudes toward the prime stimuli. Since the mid-1990s, such methods have been used to measure automatically activated attitudes toward racial or ethnic groups. For example, Fazio et al. (1995) presented pictures of Black or White male faces for a supposed recognition task interspersed with evaluative adjectives that were clearly positive or negative. The overall index of relative attitudinal differences between Black and White targets was calculated as the difference in response time (controlling for baseline responses to the various target words) across types of trials (i.e., [White/Positive – Black/Positive] – [White/Negative – Black/Negative], corresponding to the Race × Adjective Valence interaction for each respondent).

The Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) involves classifying targets (e.g., Black and White faces, representing racial categories) and attributes (e.g., love or dirt, representing the categories of pleasant and unpleasant). The IAT is an indirect measure because respondents are never asked to evaluate the targets. They simply classify any face that appears as Black or White, and they classify attribute words as either pleasant or unpleasant. Participants respond as quickly as possible using only two response options. Each corresponds to both a target and an attribute category (e.g., White and pleasant). Participants should complete the task more quickly when they associate the target with the attribute. With some methodological caveats, people generally are faster when the target category and attribute are consistent in evaluation rather than inconsistent (cf. Greenwald, Nosek, & Banaji, 2003; Nosek, Greenwald, & Banaji, 2005). Thus, relative speed in responding to items in a block with one pairing (e.g., White/ pleasant and Black/unpleasant) versus the opposite pairing (White/unpleasant and Black/pleasant) is taken as an indication of the relative differences in evaluation across the two categories of targets.

Affect Misattribution Procedure

In the affect misattribution procedure (Payne, Cheng, Govorun, & Stewart, 2005), respondents generally see a briefly presented (e.g., 75 ms) attitude object followed by a briefly presented neutral Chinese ideograph (although AMP effects are fairly robust to longer presentations of either the prime or the target; Payne et al., 2005). Respondents are not asked to evaluate the prime but instead are asked to report whether the Chinese ideograph is visually more pleasant or less pleasant than the average Chinese ideograph. Respondents should misattribute an affective response toward the attitude object (the evaluative prime) as being a reaction to the ideograph and, thus, more often report the ideograph as being visually pleasant when primed by positive rather than negative reactions. Recent research supports the role of affect misattribution per se in creating AMP responses (Gawronski & Ye, 2014).

Differences Between Explicit and Implicit Measures

Much research comparing implicit to explicit measures has shown them to be weakly correlated, if at all (especially in domains where social desirability is a concern; e.g., Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Nosek, 2005). These findings have inspired a number of research questions, including how to conceptualize the dissociations and whether the different measures predict the same or different behaviors (see also Chapter 12 in this volume).

Debate is ongoing regarding how best to understand the dissociation between implicit and explicit measures. (Compare Greenwald & Nosek, 2009, with Olson & Fazio, 2009.) Some researchers consider implicit versus explicit measures as tapping into implicit attitudes versus explicit attitudes, respectively (with the two types of attitudes represented separately in memory; e.g., Wilson, Lindsey, & Schooler, 2000, or resulting from different processes; e.g., Gawronski & Bodenhausen, 2006; Rydell, McConnell, Mackie, & Strain, 2006). Other researchers think of the measures as tapping into different aspects of the same underlying attitude representation (e.g., Fazio, 2007; Petty, Briñol, & DeMarree, 2007). Researchers favoring separable implicit and explicit attitude constructs have tended to link the different measures to the different constructs. In contrast, researchers favoring the same underlying attitude representation suggest that the measures differentially tap surrounding constructs. For example, the measures might differ in the extent to which they are affected by relevant motives to engage in certain types of responding and by opportunities to deliberate about the attitude object (e.g., Fazio, 1990; Olson & Fazio, 2009). Alternatively, the measures might differ in the extent to which they are affected by metacognitive validity assessments (e.g., Petty, Briñol, & DeMarree, 2007). For the most part, these different approaches can account for the same patterns of findings, such as implicit and explicit measures being more correlated when attitude reports are not influenced by social desirability (e.g., Nosek, 2005) or implicit measures better predicting relatively uncontrollable behaviors versus explicit measures better predicting more controllable behaviors (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Neumann, Hülsenbeck, & Seibt, 2004).

The different approaches also have inspired different research questions, however. For example, models focusing on different mental representations or processes underlying implicit versus explicit attitudes have inspired work comparing the roles of associative processes versus propositional reasoning in attitude formation and change (e.g., Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006). The MODE approach to attitude–behavior relations has focused researchers on motivation to avoid prejudiced responding as a moderator of relations between implicit and explicit measures (e.g., Fazio et al., 1995). Finally, the metacognitive model (Petty et al., 2007) has focused people on the role of perceived validity of positive and negative reactions toward the attitude object and the implications of implicit ambivalence (i.e., evaluative discrepancies between implicit and explicit reactions) for information processing (e.g., Briñol, Petty, & Wheeler, 2006).

Discrepancies between implicit and explicit measures also have invited comparisons between measures in predicting behavior. A meta-analysis by Greenwald, Poehlman, Uhlmann, and Banaji (2009) found that the IAT and explicit measures both predicted behavior, although the measures were not strongly related to one another and they tended to predict behavior in different settings. The explicit measures generally predicted behavior better than the IAT, but the IAT predicted behaviors better for attitudes potentially associated with social-desirability concerns (e.g., racial prejudice). Such conclusions were questioned by Oswald, Mitchell, Blanton, Jaccard, and Tetlock (2013), and this has resulted in continuing debate between the two research groups.

Self-Standards in Dissonance

The self-standards model of dissonance (Stone & Cooper, 2001) was aimed at reconciling traditional dissonance theory, the aversive consequences revision, the self-affirmation approach, and the self-consistency perspective. The model proposed that dissonance motivation and the role for self-related thoughts depend on the salient standards used when evaluating one's behavior. When initially evaluating their own behavior, if normative standards are salient, people compare the behavior to cultural norms. If there is a perceived discrepancy between the behavior and the norm, they experience dissonance, and that dissonance should not depend on self-views. Thus, for example, in the context of strong cultural norms for honesty, misleading a person in a dissonance study might be viewed as dishonest and make people uncomfortable regardless of how (dis)honest they typically view themselves as being (consistent with the aversive consequences approach). In contrast, if personal standards are accessible, people should compare their behavior to their personal expectations for honesty. In that case, viewing themselves more positively on that dimension should lead to greater dissonance (consistent with the self-consistency approach).

Once people experience dissonance, they try to reduce it. If people do not think about themselves when attempting to reduce dissonance, they should demonstrate typical dissonance reduction patterns unmoderated by self-esteem. However, if people think about themselves when reducing dissonance, self-esteem could enhance or reduce dissonance, depending on the relevance of the self-related standard to the dissonance-producing behavior. If people think about self-attributes relevant to the discrepant behavior, those with high self-esteem should demonstrate larger dissonance effects than those with low self-esteem (consistent with the self-consistency view). However, if people think about self-attributes that are irrelevant to the behavior (as in a typical self-affirmation task; e.g., Matz & Wood, 2005), high self-esteem should reduce dissonance effects (consistent with the self-affirmation view).

In a direct test of these ideas, Stone and Cooper (2003) asked research participants to write an uncompassionate essay. Then participants were primed either with compassion (a relevant positive self-attribute) or with creativity (an irrelevant positive self-attribute). When primed with compassion, high self-esteem led to the greatest postessay attitude change. However, when primed with creativity, low self-esteem led to the greatest postessay attitude change.

Pro-Attitudinal Versus Counterattitudinal Messages: Impact on Information Processing

As discussed in relation to the ELM, higher levels of elaboration tend to produce stronger attitudes. Therefore, an essential component of understanding attitude change is to understand the factors that influence the amount of processing of persuasive messages. It has long been posited that components of the message itself, such as whether it is pro-attitudinal or counterattitudinal, can impact the extent to which a person processes the content of the message. Indeed, classic persuasion studies suggested that counterattitudinal messages lead to people thinking more deeply (Cacioppo & Petty, 1979).

However, researchers have demonstrated instances of the exact opposite effect: pro-attitudinal messages prompting greater message processing (see Baker & Petty, 1994; Ziegler & Burger, 2011). The discrepancy motives model (Clark & Wegener, 2013) provided insight into when variables might increase or decrease scrutiny of pro- versus counterattitudinal messages. In particular, Clark and Wegener emphasized the different motives that covary with pro- versus counterattitudinal messages. Counterattitudinal messages threaten one's existing attitude. Thus, enhancing that threat through use of a credible (rather than noncredible; Clark, Wegener, Habashi, & Evans, 2012) source or a source that generally is effective rather than ineffective (Clark & Wegener, 2009) can increase processing of counterattitudinal messages. Pro-attitudinal messages do not threaten the message recipient's position directly. However, concerns about whether an ineffective (Clark & Wegener, 2009) or inexpert (Clark et al., 2012) source will successfully support the person's point of view can result in greater processing of the pro-attitudinal message than when the source is effective or expert. Along similar lines, possessing a weak premessage attitude, such as one held with ambivalence, can increase motivation to process a pro-attitudinal message (that can reduce the ambivalence; Clark, Wegener, & Fabrigar, 2008). In contrast, similar to the original notion that greater threat from a counterattitudinal message can motivate processing (e.g., Cacioppo & Petty, 1979), unambivalent premessage attitudes produce the traditional greater processing of counterattitudinal than pro-attitudinal messages.

Metacognition and Attitude Properties

One of the broadest, and most essential, advances in recent research has been an increased focus on metacognition. Metacognition research has made its mark in virtually every area of attitude research. As such, the next subsections demonstrate developments in a few attitude domains rather than serve as an exhaustive list.

Embodied Cognition

The relations among conceptual, perceptual, and motor processes have garnered significant interest in cognitive and social psychology (see Barsalou, 1999; Mahon & Caramazza, 2008; Petty, Briñol, Tormala, & Wegener, 2007; Wilson, 2002). Connections between emotional or evaluative experience and physical movement have become known as embodied cognition. In the domain of attitudes, Wells and Petty (1980) conducted one of the earliest studies of embodied cognition. They had participants nod their head or shake their head while listening to persuasive information. They found that receipt of agreeable information led to faster head-nodding but slower head-shaking, and receipt of disagreeable information led to faster head shaking but slower head nodding. Briñol and Petty (2003) further demonstrated that nodding versus shaking had different effects on persuasion, depending on the level of argument quality. When a persuasive message's arguments were strong, head-nodding produced greater persuasion than head-shaking. However, when the persuasive message arguments were weak, the opposite effect occurred, and head-nodding reduced persuasion. This study was understood within the broader self-validation hypothesis, in which the head movements affected people's confidence in the thoughts they were having at the time (see Petty, Briñol, & Tormala, 2002).

Self-Validation

Whereas attitude certainty has been in the literature for some time, work on validation of thoughts has come on the scene more recently. Within the ELM, validation of one's own thoughts serves as an additional role for persuasion variables. The self-validation hypothesis proposes that generating thoughts in response to a message does not necessarily mean those thoughts will influence attitudes. Rather, thoughts have an influence to the extent one is confident in those thoughts (Petty et al., 2002).

Validation effects are most likely to occur with high levels of elaboration because people need to be generating thoughts for a variable to validate those thoughts. Also, the validating variable generally comes after the message, so thoughts are in mind at the time of the validating experience. A number of variables can affect the extent to which people feel confident in their thoughts. For example, people who are made to feel powerful or happy after receiving a message use their thoughts more than people made to feel powerless or sad (Briñol, Petty, & Barden, 2007; Briñol, Petty, Valle, Rucker, & Becerra, 2007; see Briñol & Petty, 2009, for a review).

Perceived Elaboration

As noted earlier, elaboration is thought to be a key determinant of the strength of the resulting attitude (Petty et al., 1995). Although this link between elaboration and strength has been well established, only recent research has examined the role of perceptions of elaboration in these effects. In particular, Barden and Petty (2008) argued that elaboration is often conscious, in that people can report on the amount of processing in which they have engaged. Barden and Petty suggested that perceived elaboration is associated via a thoughtfulness heuristic with greater attitude certainty. The resulting certainty may then be responsible for at least some of the strength-related outcomes of elaboration. Through a series of studies, Barden and Petty demonstrated that traditional antecedents of message-relevant thinking (e.g., need for cognition and distraction) impact perceptions of elaboration, and perceptions of elaboration link to attitude certainty. Further studies demonstrated that manipulation of perceived elaboration also influences certainty, independent of the actual amount of thought in which participants engaged.

Subsequent studies have continued to explore the factor of perceived elaboration and have identified cases in which effort on alternative activities bleed into perceptions of effort in elaborating. Wan, Rucker, Tormala, and Clarkson (2010) found that regulatory depletion tasks (e.g., crossing out the letter *e* each time it appears, following a strict set of rules) can impact perceptions of elaboration: The more depleted participants were, the more perceived elaboration they had for a subsequent task and, subsequently, the more certain they were in their attitudes.

Metabases Versus Structural Bases

One last area of recent development has focused on perceptions of the bases of attitudes (known as metabases; see, Petty & Fabrigar, 2013). The research on metabases builds on traditional attitudes work suggesting that the structural foundation of an attitude may be based primarily on cognitive or affective information (e.g., Crites, Fabrigar, & Petty, 1994). The perceived foundation of attitudes has been shown to uniquely impact interest in affective or cognitive information beyond the directly measured structure of the attitude (See, Petty, & Fabrigar, 2008). Moreover, recent research suggests that metabases better predict processing interest whereas structural bases better predict processing efficiency (See et al., 2013).

FUTURE DIRECTIONS

It may be somewhat hazardous to predict future directions a literature may take. A safe bet is that many of the recent trends will continue, as they represent active and ongoing programs of research. In that regard, we anticipate much ongoing work examining implicit/automatic processes for attitude change and for change on implicit versus explicit measures of attitudes (which then may predict different kinds of behaviors or behavior in different settings). This work is likely to build on existing models of attitude representation in memory and on process assumptions underlying responses to the various implicit and explicit measures.

Likewise, research on dissonance is likely to continue. With recent challenges to the role of dissonance in spreading of alternatives following free choice (e.g., Chen & Risen, 2010), efforts documenting dissonance-based discomfort are likely. Also, there has been no definitive solution to whether cognitive dissonance requires or involves effortful thinking (cf. Jarcho, Berkman, & Lieberman, 2011 and Lieberman, Ochsner, Gilbert, & Schacter, 2001 with Gawronski & Strack, 2004) or when it does or does not. Thus, we would expect these issues to receive ongoing attention.

In the area of attitude–behavior consistency, despite over 40 years of research identifying moderators of such consistency, many interesting questions remain. For example, Fabrigar, Wegener, and MacDonald (2010) noted that many studies leave open questions regarding whether observed differences in attitude–behavior relations stem from differential influences of attitudes at the time of behavior or from mere differences in prediction of behaviors over time. Prediction refers to how well an attitude measure corresponds with a behavioral outcome. For an attitude to predict a behavior, the measure must capture the attitude accurately. Additionally, the attitude must be stable over time. Thus, one could have differences in behavior prediction by attitudes even if attitudes at the time of the behavior are having equal influence on those behaviors (either because accuracy of the attitude measure differed or because attitude stability differed across levels of a moderating variable). Similar to some elements of both the MODE and ELM frameworks, differences in influence of attitudes on behaviors might occur at low or high levels of deliberation. At low levels of thought about the behavior, attitudes may serve as direct or indirect cues informing people whether they should engage in a behavior. At higher levels of thought about the behavior, people's attitudes could serve as a reason to engage in a particular behavior or could influence the direction of thoughts that people generate about a behavior.

In areas related to attitude strength, many open questions remain about how the various properties of attitudes relate to one another and moderate each other's effects. For example, repeated expression (a typical manipulation of attitude accessibility) increases attitude certainty (Holland, Verplanken, & van Knippenberg, 2003; Petrocelli, Tormala, & Rucker, 2007). A number of other strength-related properties of attitudes may serve as antecedents to other properties, such that the later properties mediate effects on behavioral or information processing outcomes. Alternatively, the different attitude properties also might interact in interesting ways. For example, when attitudes are held with both certainty and ambivalence, the attitudes function as if they are more ambivalent than when they are held with uncertainty and ambivalence (e.g., Clarkson, Tormala, & Rucker, 2008; Luttrell, Petty, & Briñol, 2016). Relatedly, recent work on bolstering effects have identified cases in which attitudes associated with doubt or ambivalence (traditionally weak attitudes) have been associated with more attitude-consistent choices than attitudes associated with confidence or univalence (e.g., Sawicki et al., 2011, 2013). Such findings suggest that there must be important moderators of when traditional strength effects occur and when they actually might be reversed.

Finally, the potential role of contingency awareness and propositional reasoning in evaluative conditioning remains an active point of contention (e.g., see Gawronski & Walther, 2012; Pleyers, Corneille, Luminet, & Yzerbyt, 2007). However, evaluative conditioning effects have been found when the conditioned stimulus, unconditioned stimulus, or both are presented subliminally (e.g., De Houwer, Baeyens, & Eelen, 1994; Dijksterhuis, 2004). We suspect that, like cognitive dissonance and attitude change more generally, a multiprocess view is likely to be useful in determining how and when different proposed mechanisms play a role in evaluative conditioning (see Jones, Olson, & Fazio, 2010, for discussion).

CONCLUSION

In this chapter, we began by defining attitudes as summary evaluations and describing both direct and indirect methods of measuring them. Next we examined when attitudes predict behavior. Despite early contentions that attitudes are poor predictors of behavior, research has identified a number of factors that help determine when strong rather than weak relations are found between attitudes and behavior. We also discussed research on how to change attitudes. Many separate theories have been developed over time, but meta-theoretical approaches have helped to organize the previous efforts. In the research on dissonance, the self-standards model organizes when self-esteem and other self-related views do or do not play a role in cognitive inconsistency leading to changes in attitudes. In attitude change more generally, the elaboration likelihood model takes an even broader approach. By arraying previous theories and mechanisms along an elaboration continuum, the ELM allows one to make predictions for limiting conditions on previous effects.

Also, it makes specific predictions for the kinds of mechanisms likely responsible for effects of persuasion variables across that elaboration continuum. Such developments helped get the attitudes and persuasion literature past various criticisms raised in the 1970s and facilitated a new vigor and momentum that has continued to this day. It is our hope that some of the lessons learned in this literature can be more broadly helpful to social psychology as various crises have been raised recently about replicability of results. Perhaps the current crises, like those of the 1970s, will prove to be opportunities for theoretical development.

REFERENCES

Abelson, R. P. (1959). Modes of resolution of belief dilemmas. *Journal of Conflict Resolution*, *3*, 343–352.

Abelson, R. P. (1972). Are attitudes necessary? In B. T. King & E. McGinnies (Eds.), *Attitudes, Conflict, and Social Change* (pp. 19–32). New York, NY: Academic Press.

Ajzen, I. (1985). From intentions to actions: a theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action Control: From Cognitions to Behaviors* (pp. 11–39). New York, NY: Springer.

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: a theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918.

Albarracín, D., & Wyer, R. S. Jr. (2000). The cognitive impact of past behavior: influences on beliefs, attitudes, and future behavioral decisions. *Journal of Personality and Social Psychology*, *79*(1), 5–22.

Armitage, C. J., & Conner, M. (2000). Attitudinal ambivalence: a test of three key hypotheses. *Personality and Social Psychology Bulletin*, *26*(11), 1421–1432.

Aronson, E. (1969). The theory of cognitive dissonance: a current perspective. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 4, pp. 1–34). San Diego, CA: Academic Press.

Aronson, E., & Carlsmith, J. M. (1962). Performance expectancy as a determinant of actual performance. *Journal of Abnormal and Social Psychology*, *65*(3), 178–182.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, *59*(2), 177–181.

Baeyens, F., Eelen, P., & Van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: a case for unaware affective-evaluative learning. *Cognition and Emotion*, *4*(1), 3–18.

Baker, S. M., & Petty, R. E. (1994). Majority and minority influence: source-position imbalance as a determinant of message scrutiny. *Journal of Personality and Social Psychology*, *67*(1), 5–19.

Barden, J., & Petty, R. E. (2008). The mere perception of elaboration creates attitude certainty: exploring the thoughtfulness heuristic. *Journal of Personality and Social Psychology*, *95*(3), 489–509.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*(1), 104–128.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, *22*(04), 637–660.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: the case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*(4), 637–653.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.

Boniecki, K. A., & Jacks, J. Z. (2002). The elusive relationship between measures of implicit and explicit prejudice. *Representative Research in Social Psychology*, *26*, 1–14.

Boninger, D. S., Krosnick, J. A., Berent, M. K., & Fabrigar, L. R. (1995). The causes and consequences of attitude importance. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength:*

Antecedents and Consequences (pp. 159–189). Mahwah, NJ: Erlbaum.

Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, *63*, 545–552.

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology*, *52*(3), 384–389.

Briñol, P., & Petty, R. E. (2003). Overt head movements and persuasion: a self-validation analysis. *Journal of Personality and Social Psychology*, *84*(6), 1123–1139.

Briñol, P., & Petty, R. E. (2009). Persuasion: insights from the self-validation hypothesis. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 41, pp. 69–118). New York, NY: Elsevier.

Briñol, P., & Petty, R. E. (2012). A history of attitudes and persuasion research. In A. W. Kruglansi & W. Stroebe (Eds.), *Handbook of the History of Social Psychology* (pp. 283–318). New York, NY: Psychology Press.

Briñol, P., Petty, R. E., & Barden, J. (2007). Happiness versus sadness as determinants of thought confidence in persuasion: a self-validation analysis. *Journal of Personality and Social Psychology*, *93*, 711–727.

Briñol, P., Petty, R. E., Valle, C., Rucker, D. D., & Becerra, A. (2007). The effects of message recipients' power before and after persuasion: a self-validation analysis. *Journal of Personality and Social Psychology*, *93*, 1040–1053.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: consequences for information processing. *Journal of Personality and Social Psychology*, *91*, 154–170.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, *47*(6), 1191–1205.

Byrne, D. E. (1971). *The Attraction Paradigm*. New York, NY: Academic Press.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: a critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*(3), 401–423.

Cacioppo, J. T., & Petty, R. E. (1979). Effects of message repetition and position on cognitive response, recall, and persuasion. *Journal of Personality and Social Psychology*, *37*(1), 97–109.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: the life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Cacioppo, J. T., Petty, R. E., Kao, C. F., & Rodriguez, R. (1986). Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology*, *51*, 1032–1043.

Calder, B. J., Ross, M., & Insko, C. A. (1973). Attitude change and attitude attribution: effects of incentive, choice, and consequences. *Journal of Personality and Social Psychology*, *25*, 84–99.

Carpenter, C. J. (2012). A meta-analysis of the functional matching effect based on functional attitude theory. *Southern Communication Journal*, *77*(5), 438–451.

Cervellon, M. C., & Dubé, L. (2002). Assessing the cross-cultural applicability of affective and cognitive components of attitude. *Journal of Cross-Cultural Psychology*, *33*(3), 346–357.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752–766.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (pp. 212–252). New York, NY: Guilford Press.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgment.

Journal of Personality and Social Psychology, *66*(3), 460–473.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (pp. 73–96). New York, NY: Guilford Press.

Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: revisiting the free-choice paradigm. *Journal of Personality and Social Psychology*, *99*(4), 573–594.

Clark, J. K., Evans, A. T., & Wegener, D. T. (2011). Perceptions of source efficacy and persuasion: multiple mechanisms for source effects on attitudes. *European Journal of Social Psychology*, *41*(5), 596–607.

Clark, J. K., & Wegener, D. T. (2009). Source entitativity and the elaboration of persuasive messages: the roles of perceived efficacy and message discrepancy. *Journal of Personality and Social Psychology*, *97*(1), 42–57.

Clark, J. K., & Wegener, D. T. (2013). Message position, information processing, and persuasion: the Discrepancy Motives Model. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 189–232). San Diego, CA: Academic Press.

Clark, J. K., Wegener, D. T., & Fabrigar, L. R. (2008). Attitude ambivalence and message-based persuasion: motivated processing of proattitudinal information and avoidance of counterattitudinal information. *Personality and Social Psychology Bulletin*, *34*, 565–577.

Clark, J. K., Wegener, D. T., Habashi, M. M., & Evans, A. T. (2012). Source expertise and persuasion: the effects of perceived opposition or support on message scrutiny. *Personality and Social Psychology Bulletin*, *38*(1), 90–100.

Clarkson, J. J., Tormala, Z. L., & Rucker, D. D. (2008). A new look at the consequences of attitude certainty: the amplification hypothesis. *Journal of Personality and Social Psychology*, *95*(4), 810–825.

Clarkson, J. J., Tormala, Z. L., & Leone, C. (2011). A self-validation perspective on the mere thought effect. *Journal of Experimental Social Psychology*, *47*(2), 449–454.

Collins, B. E., & Hoyt, M. F. (1972). Personal responsibility-for-consequences: an integration and extension of the "forced compliance" literature. *Journal of Experimental Social Psychology*, *8*(6), 558–593.

Cook, T. D., & Flay, B. R. (1978). The persistence of experimentally induced attitude change. *Advances in Experimental Social Psychology*, *11*, 1–57.

Cooper, J., & Duncan, B. L. (1971). Cognitive dissonance as a function of self-esteem and logical inconsistency. *Journal of Personality*, *39*(2), 289–302.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 17, pp. 229–266). New York, NY: Academic Press.

Crites, S. L., Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1995). Bioelectrical echoes from evaluative categorization: II. A late positive brain potential that varies as a function of attitude registration rather than attitude report. *Journal of Personality and Social Psychology*, *68*(6), 997–1013.

Crites, S. L., Fabrigar, L. R., & Petty, R. E. (1994). Measuring the affective and cognitive properties of attitudes: conceptual and methodological issues. *Personality and Social Psychology Bulletin*, *20*(6), 619–634.

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: evaluative processing goals shape amygdala activity. *Psychological Science*, *19*(2), 152–160.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: a multilevel framework for attitudes and evaluation. *Social Cognition*, *25*(5), 736–760.

Davidson, A. R., & Jaccard, J. J. (1979). Variables that moderate the attitude–behavior relation: results of a longitudinal survey. *Journal of Personality and Social Psychology*, *37*(8), 1364–1376.

DeBono, K. G., & Packer, M. (1988). "The effects of image and quality based advertising on perceptions of product quality." Unpublished manuscript, Union College, Schenectady, NY.

De Houwer, J., Baeyens, F., & Eelen, P. (1994). Verbal evaluative conditioning with undetected US presentations. *Behaviour Research and Therapy*, *32*(6), 629–633.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychological Bulletin*, *135*(3), 347–368.

Dijksterhuis, A. (2004). I like myself but I don't know why: enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology*, *86*, 345–355.

Doll, J., & Ajzen, I. (1992). Accessibility and stability of predictors in the theory of planned behavior. *Journal of Personality and Social Psychology*, *63*(5), 754–765.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*(5), 510–540.

Eagly, A. H., & Chaiken, S. (1993). *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Eagly, A. H., & Chaiken, S. (1998). Attitude structure. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology* (pp. 269–322). New York, NY: McGraw-Hill.

Eagly, A. H., Kulesa, P., Brannon, L. A., Shaw, K., & Hutson-Comeaux, S. (2000). Why counterattitudinal messages are as memorable as proattitudinal messages: the importance of active defense against attack. *Personality and Social Psychology Bulletin*, *26*(11), 1392–1408.

Elkin, R. A., & Leippe, M. R. (1986). Physiological arousal, dissonance, and attitude change: evidence for a dissonance-arousal link and a "Don't remind me" effect. *Journal of Personality and Social Psychology*, *51*, 55–65.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, *67*(3), 382–394.

Fabrigar, L. R., MacDonald, T., & Wegener, D. T. (2005). The structure of attitudes. In D. Albarracín, B. Johnson, & M. Zanna (Eds.), *The Handbook of Attitudes* (pp. 79–124). Mahwah, NJ: Erlbaum.

Fabrigar, L. R., & Petty, R. E. (1999). The role of the affective and cognitive bases of attitudes in susceptibility to affectively and cognitively based persuasion. *Personality and Social Psychology Bulletin*, *25*(3), 363–381.

Fabrigar, L. R., Wegener, D. T., & MacDonald, T. K. (2010). Distinguishing between prediction and influence: multiple processes underlying attitude–behavior consistency. In C. R. Agnew, D. E. Carlston, W. G. Graziano, & J. R. Kelly (Eds.), *Then a Miracle Occurs: Focusing on Behavior in Social Psychological Theory and Research* (pp. 162–185). New York, NY: Oxford University Press.

Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (pp. 204–243). New York, NY: Guilford Press.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: the MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 75–109). San Diego, CA: Academic Press.

Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: data reanalysis and commentary on Bargh, Chaiken, Govender, and Pratto (1992). *Journal of Personality and Social Psychology*, *64*(5), 753–758.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 247–282). Hillsdale, NJ: Erlbaum.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, *25*(5), 603–637.

Fazio, R. H., Blascovich, J., & Driscoll, D. M. (1992). On the functional value of attitudes: the influence of accessible attitudes on the ease and quality of decision making. *Personality and Social Psychology Bulletin*, *18*(4), 388–401.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013–1027.

Fazio, R. H., Lenn, T. M., & Effrein, E. A. (1984). Spontaneous attitude formation. *Social Cognition*, *2*(3), 217–234.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude–perception and attitude–behavior relations: an investigation of the 1984 presidential election. *Journal of Personality and Social Psychology*, *51*(3), 505–514.

Fazio, R. H., Zanna, M. P., & Cooper, J. (1977). Dissonance and self-perception: an integrative view of each theory's proper domain of application. *Journal of Experimental Social Psychology*, *13*(5), 464–479.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Festinger, L. (Ed.) (1964). *Conflict, Decision, and Dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*(2), 203–210.

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Freedman, J. L. (1965). Long-term behavioral effects of cognitive dissonance. *Journal of Experimental Social Psychology*, *1*(2), 145–155.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, *15*(3), 485–499.

Gawronski, B., & Payne, B. K. (Eds.). (2011). *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. New York, NY: Guilford Press.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: dissonance changes explicit but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*(4), 535–542.

Gawronski, B., & Walther, E. (2012). What do memory data tell us about the role of contingency awareness in evaluative conditioning? *Journal of Experimental Social Psychology*, *48*(3), 617–623.

Gawronski, B., & Ye, Y. (2014). What drives priming effects in the affect misattribution procedure? *Personality and Social Psychology Bulletin*, *40*(1), 3–15.

Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: a meta-analysis of the attitude–behavior relation. *Psychological Bulletin*, *132*(5), 778–822.

Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological Foundations of Attitudes* (pp. 147–170). New York, NY:

Greenwald, A. G. (1975). On the inconclusiveness of "crucial" cognitive tests of dissonance versus self-perception theories. *Journal of Experimental Social Psychology*, *11*(5), 490–499.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Greenwald, A. G., & Nosek, B. A. (2009). Attitudinal dissociation: what does it mean? In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the New Implicit Measures* (pp. 65–82). New York, NY: Psychology Press.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41.

Hammond, K. R. (1948). Measuring attitudes by error-choice: an indirect method. *Journal of Abnormal and Social Psychology*, *43*(1), 38–48.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: a case study. *Journal of Abnormal and Social Psychology*, *49*(1), 129–134.

Haugtvedt, C. P., & Petty, R. E. (1992). Personality and persuasion: need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology*, *63*, 308–319.

Haugtvedt, C. P., & Wegener, D. T. (1994). Message order effects in persuasion: an attitude strength perspective. *Journal of Consumer Research*, *21*, 205–218.

Heesacker, M., Petty, R. E., & Cacioppo, J. T. (1983). Field dependence and attitude change: source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality*, *51*(4), 653–666.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. Hillsdale, NJ: Erlbaum.

Hildum, D. C., & Brown, R. W. (1956). Verbal reinforcement and interviewer bias. *Journal of Abnormal and Social Psychology*, *53*(1), 108–111

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*(10), 1369–1385.

Holland, R. W., Verplanken, B., & van Knippenberg, A. (2003). From repetition to conviction: attitude accessibility as a determinant of attitude certainty. *Journal of Experimental Social Psychology*, *39*(6), 594–601.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: making objective judgments subjectively. *Social Cognition*, *7*(1), 51–66.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and Persuasion: Psychological Studies of Opinion Change*. New Haven, CT: Yale University Press.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). Experiments on mass communication. In *Studies in Social Psychology in World War II* (Vol. 3). Princeton, NJ: Princeton University Press.

Insko, C. A. (1964). Primacy versus recency in persuasion as a function of the timing of arguments and measures. *Journal of Abnormal and Social Psychology*, *69*, 381–391.

Insko, C. A. (1965). Verbal reinforcement of attitude. *Journal of Personality and Social Psychology*, *2*, 621–623.

James, W. (1890/1950). *The Principles of Psychology* (Vol. 1). New York, NY: Dover.

Janis, I. L., & King, B. T. (1954). The influence of role playing on opinion change. *Journal of Abnormal and Social Psychology*, *49*, 211–218.

Jarcho, J. M., Berkman, E. T., & Lieberman, M. D. (2011). The neural basis of rationalization: cognitive dissonance reduction during decision-making. *Social Cognitive and Affective Neuroscience*, *6*(4), 460–467.

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology*, *70*(1), 172–194.

Jones, C. R., Olson, M. A., & Fazio, R. H. (2010). Evaluative conditioning: the "how" question. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 43, pp. 205–255). San Diego, CA: Academic Press.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*(2), 163–204.

Katz, D., Sarnoff, I., & McClintock, C. (1956). Ego-defense and attitude change. *Human Relations*, *9*, 27–46.

Kiesler, S. B., & Mathog, R. (1968). The distraction hypothesis in attitude change. *Psychological Reports*, *23*, 1123–1133.

Krosnick, J. A. (1988). Attitude importance and attitude change. *Journal of Experimental Social Psychology*, *24*(3), 240–255.

Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K., & Carnot, C. G. (1993). Attitude strength: one construct or many related constructs? *Journal of Personality and Social Psychology*, *65*(6), 1132–1151.

Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of attitudes. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The Handbook of Attitudes* (pp. 21–76). Mahwah, NJ: Erlbaum.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: an overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 1–24). Mahwah, NJ: Erlbaum.

Kruglanski, A. W., Freund, T., & Bar-Tal, D. (1996). Motivational effects in the mere-exposure paradigm. *European Journal of Social Psychology*, *26*(3), 479–499.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, *207*(4430), 557–558.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776–785.

Lemon, N. (1973). *Attitudes and Their Measurement*. Oxford, UK: Wiley.

Liberzon, I., Phan, K. L., Decker, L. R., & Taylor, S. F. (2003). Extended amygdala and emotional salience: a PET activation study of positive and negative affect. *Neuropsychopharmacology*, *28*, 726–733.

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, *12*(2), 135–140.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *140*, 1–55.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098–2109.

Losch, M. E., & Cacioppo, J. T. (1990). Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology*, *26*, 289–304.

Luttrell, A., Petty, R. E., & Briñol, P. (2016). Attitude ambivalence and certainty can interact to predict attitude stability over time. *Journal of Experimental Social Psychology*, *63*, 56–68.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*(5), 808–817.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology* (Paris), *102*(1), 59–70.

Martin, L. L., Abend, T., Sedikides, C., & Green, J. D. (1997). How would it feel if . . . ? Mood as input to a role fulfillment evaluation process. *Journal of Personality and Social Psychology*, *73*(2), 242–253.

Matz, D. C., & Wood, W. (2005). Cognitive dissonance in groups: the consequences of disagreement. *Journal of Personality and Social Psychology*, *88*(1), 22–37.

McGuire, W. J. (1968). Personality and attitude change: an information-processing theory. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Ed.), *Psychological Foundations of Attitudes* (pp. 171–196). New York, NY: Academic Press.

Michaelis, W., & Eysenck, H. J. (1971). The determination of personality inventory factor patterns and intercorrelations by changes in real-life motivation. *Journal of Genetic Psychology*, *118*(2), 223–234.

Miller, N., & Campbell, D. T. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *Journal of Abnormal and Social Psychology*, *59*, 1–9.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: specific, general, and diffuse effects. *Psychological Science*, *11*, 462–466.

Moreland, R. L., & Beach, S. R. (1992). Exposure effects in the classroom: the development of affinity among students. *Journal of Experimental Social Psychology*, *28*(3), 255–276.

Munro, G. D., & Ditto, P. H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin*, *23*(6), 636–653.

Neumann, R., Hülsenbeck, K., & Seibt, B. (2004). Attitudes towards people with AIDS and avoidance behavior: automatic and reflective bases of behavior. *Journal of Experimental Social Psychology*, *40*(4), 543–550.

Newcomb, T. M. (1943). *Personality and Social Change: Attitude Formation in a Student Community*. New York, NY: Dryden.

Nier, J. A. (2005). How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach. *Group Processes & Intergroup Relations*, *8*(1), 39–52.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*(4), 565–584.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, *31*(2), 166–180.

Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: the perspective of the MODE model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.). *Attitudes: Insights from the New Implicit Measures* (pp. 19–63). New York, NY: Psychology Press.

Osgood, C.E., Suci, G., & Tannenbaum, P. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*(1), 42–55.

Ostrom, T. M. (1969). The relationship between the affective, behavioral, and cognitive components of attitude. *Journal of Experimental Social Psychology*, *5*(1), 12–30.

Ostrom, T. M. (1989). Interdependence of attitude theory and measurement. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude Structure and Function* (pp. 11–36). Hillsdale, NJ: Erlbaum.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171–192.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277–293.

Pelham, B. W. (1991). On confidence and consequence: the certainty and importance of self-knowledge. *Journal of Personality and Social Psychology*, *60*(4), 518–530.

Petrocelli, J. V., Tormala, Z. L., & Rucker, D. D. (2007). Unpacking attitude certainty: attitude clarity and attitude correctness. *Journal of Personality and Social Psychology*, *92*(1), 30–41.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The meta-cognitive model (MCM) of attitudes: implications for attitude measurement, change, and strength. *Social Cognition*, *25*(5), 657–686.

Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: the self-validation hypothesis. *Journal of Personality and Social Psychology*, *82*(5), 722–741.

Petty, R. E., Briñol, P., Tormala, Z. L., & Wegener, D. T. (2007). The role of meta-cognition in social judgment. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 254–284). New York, NY: Guilford Press.

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, *37*(10), 1915–1926.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (pp. 123–205). New York, NY: Academic Press.

Petty, R. E., & Cacioppo, J. T. (1996). *Attitudes and Persuasion: Classic and Contemporary Approaches*. Boulder, CO: Westview Press.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, *41*, 847–855.

Petty, R. E., Fazio, R. H., & Briñol, P (Eds.) (2009). *Attitudes: Insights from the New Implicit Measures*. New York, NY: Psychology Press.

Petty, R. E., Haugtvedt, C. P., & Smith, S. M. (1995). Elaboration as a determinant of attitude strength. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 93–130). Mahwah, NJ: Erlbaum.

Petty, R. E., Schumann, D. W., Richman, S. A., & Strathman, A. J. (1993). Positive mood and persuasion: different roles for affect under high- and low-elaboration conditions. *Journal of Personality and Social Psychology*, *64*, 5–20.

Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology*, *48*, 609–647.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: multiple roles for persuasion variables. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.). *The Handbook of Social Psychology* (pp. 323–390). New York, NY: McGraw-Hill.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology* (pp. 41–72). New York, NY: Guilford Press.

Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: thought disruption versus effort justification. *Journal of Personality and Social Psychology*, *34*(5), 874–884.

Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis)liking: item-based analyses reveal that valence acquisition via evaluative conditioning emerges only when there is contingency awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 130–144.

Preston, S. D., & De Waal, F. (2002). Empathy: its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20.

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology*, *71*(3), 431–449.

Priester, J. R., & Petty, R. E. (2001). Extending the bases of subjective attitudinal ambivalence: interpersonal and intrapersonal antecedents of evaluative tension. *Journal of Personality and Social Psychology*, *80*(1), 19–34.

Regan, D. T., & Fazio, R. (1977). On the consistency between attitudes and behavior: look to the method of attitude formation. *Journal of Experimental Social Psychology*, *13*(1), 28–45.

Rhodes, N., & Wood, W. (1992). Self-esteem and intelligence affect influenceability: the mediating role of message reception. *Psychological Bulletin*, *111*(1), 156–171.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude Organization and Change: An Analysis of Consistency among Attitude Components* (pp. 1–14). New Haven, CT: Yale University Press.

Roskos-Ewoldsen, D. R., & Fazio, R. H. (1992). On the orienting value of attitudes: attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology*, *63*(2), 198–211.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*(6), 995–1008.

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science*, *17*(11), 954–958.

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality and Social Psychology*, *59*(4), 614–622.

Sawicki, V., Wegener, D. T., Clark, J. K., Fabrigar, L. R., Smith, S. M., & Bengal, S. T. (2011). Seeking confirmation in times of doubt: selective exposure and the motivational strength of weak attitudes. *Social Psychological and Personality Science*, *2*(5), 540–546.

Sawicki, V., Wegener, D. T., Clark, J. K., Fabrigar, L. R., Smith, S. M., & Durso, G. R. O.

(2013). Feeling conflicted and seeking information: when ambivalence enhances and diminishes selective exposure to attitude-consistent information. *Personality and Social Psychology Bulletin*, *39*, 735–747.

Schwarz, N. (1990). Feelings as information: informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (Vol. 2, pp. 527–561). New York, NY: Guilford Press.

Schwarz, N. (1999). Self-reports: how the questions shape the answers. *American Psychologist*, *54*(2), 93–105.

Schuette, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin*, *21*(7), 704–710.

See, Y. H. M., Petty, R. E., & Fabrigar, L. R. (2008). Affective and cognitive meta-bases of attitudes: unique effects on information interest and persuasion. *Journal of Personality and Social Psychology*, *94*(6), 938.

See, Y. H. M., Petty, R. E., & Fabrigar, L. R. (2013). Affective–cognitive meta-bases versus structural bases of attitudes predict processing interest versus efficiency. *Personality and Social Psychology Bulletin*, *39*(8), 1111–1123.

Shavitt, S. (1990). The role of attitude objects in attitude functions. *Journal of Experimental Social Psychology*, *26*(2), 124–148.

Shavitt, S., Swan, S., Lowrey, T. M., & Wänke, M. (1994). The interaction of endorser attractiveness and involvement in persuasion depends on the goal that guides message processing. *Journal of Consumer Psychology*, *3*(2), 137–162.

Sherif, M., & Hovland, C. I. (1961). *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven, CT: Yale University Press.

Sherif, M. (1977). Crisis in social psychology: some remarks towards breaking through the crisis. *Personality and Social Psychology Bulletin*, *3*(3), 368–382.

Sigall, H. & Page, R. (1971). Current stereotypes: a little fading, a little faking. *Journal of Personality and Social Psychology*, *18*, 247–255.

Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: the forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology*, *68*(2), 247–260.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 397–420). Cambridge, UK: Cambridge University Press.

Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology*, *71*(5), 888–898.

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and Personality*. New York, NY: Wiley.

Smith, S. M., Fabrigar, L. R., Powell, D. M., & Estrada, M.-J. (2007). The role of information processing capacity and goals in attitude-congruent selective exposure effects. *Personality and Social Psychology Bulletin*, *33*, 948–960.

Snyder, M., & DeBono, K. G. (1985). Appeals to image and claims about quality: understanding the psychology of advertising. *Journal of Personality and Social Psychology*, *49*(3), 586–597.

Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology*, *57*(1), 37–40.

Steele, C. M. (1988). The psychology of self-affirmation: sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 21, pp. 261–302). New York, NY: Academic Press.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*(1), 5–19.

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: the role of affirmational resources. *Journal of Personality and Social Psychology*, *64*(6), 885–896.

Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology*, *37*, 228–243.

Stone, J., & Cooper, J. (2003). The effect of self-attribute relevance on how self-esteem moderates attitude change in dissonance processes. *Journal of Experimental Social Psychology*, *39*(5), 508–515.

Tesser, A. (1976). Attitude polarization as a function of thought and reality constraints. *Journal of Research in Personality*, *10*(2), 183–194.

Thompson, M. M., Zanna, M. P., & Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 361–386). New York, NY: Psychology Press.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, *33*(4), 529–554.

Thurstone, L. L., & Chave, E. J. (1929). *The Measurement of Attitude*. Chicago: University of Chicago Press.

Tormala, Z. L., & Rucker, D. D. (2007). Attitude certainty: a review of past findings and emerging perspectives. *Social and Personality Psychology Compass*, *1*(1), 469–492.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, *103*(3), 299–314.

Trafimow, D., & Finlay, K. A. (1996). The importance of subjective norms for a minority of people: between-subjects and within-subjects analyses. *Personality and Social Psychology Bulletin*, *22*(8), 820–828.

Van Bavel, J. J., Xiao, Y., & Cunningham, W. A. (2012). Evaluation is a dynamic process: moving beyond dual system models. *Social and Personality Psychology Compass*, *6*(6), 438–454.

Wan, E. W., Rucker, D. D., Tormala, Z. L., & Clarkson, J. J. (2010). The effect of regulatory depletion on attitude certainty. *Journal of Marketing Research*, *47*(3), 531–541.

Watson, J. B. (1925). *Behaviorism*. New York, NY: People's Institute.

Watts, W. A. (1967). Relative persistence of opinion change induced by active compared to passive participation. *Journal of Personality and Social Psychology*, *5*, 4–15.

Wegener, D. T., & Carlston, D. E. (2005). Cognitive processes in attitude formation and change. In D. Albarracín, B. Johnson, & M. Zanna (Eds.), *The Handbook of Attitudes* (pp. 493–542). Mahwah, NJ: Erlbaum.

Wegener, D. T., & Fabrigar, L. R. (2004). Constructing and evaluating quantitative measures for social psychological research: conceptual challenges and methodological solutions. In C. Sansone, C. C. C. Morf, & A. T. Panter (Eds.), *The SAGE Handbook of Methods in Social Psychology* (pp. 145–172). New York, NY: SAGE.

Wegener, D. T., & Petty, R. E. (2013). Attitudes and social cognition as social psychological siblings. In D. Carlston (Ed.), *Oxford Handbook of Social Cognition* (pp. 118–143). New York, NY: Oxford University Press.

Wegener, D. T., Petty, R. E., & Klein, D. J. (1994). Effects of mood on high elaboration attitude change: the mediating role of likelihood judgments. *European Journal of Social Psychology*, *24*, 25–44.

Wegener, D. T., Petty, R. E., & Smith, S. M. (1995). Positive mood can increase or decrease message scrutiny: the hedonic contingency view of mood and message processing. *Journal of Personality and Social Psychology*, *69*, 5–15.

Wells, G. L., & Petty, R. E. (1980). The effects of overt head movements on persuasion: compatibility and incompatibility of responses. *Basic and Applied Social Psychology*, *1*(3), 219–230.

Westie, F. R. (1953). A technique for the measurement of race attitudes. *American Sociological Review*, *18*(1), 73–78.

Wicker, A. W. (1969). Attitudes versus actions: the relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, *25*(4), 41–78.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*(1), 101–126.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*(6), 989–1000.

Wood, W. (1982). Retrieval of attitude-relevant information from memory: effects on susceptibility to persuasion and on intrinsic motivation. *Journal of Personality and Social Psychology*, *42*(5), 798–810.

Wood, W., Rhodes, N. D., & Biek, M. (1995). Working knowledge and attitude strength: an information-processing analysis. In R. Petty & J. Krosnick (Eds.), *Attitude Strength: Antecedents and Consequences* (pp. 283–313). Hillsdale, NJ: Erlbaum.

Ybarra, O., & Trafimow, D. (1998). How priming the private self or collective self affects the relative weights of attitudes and subjective norms. *Personality and Social Psychology Bulletin*, *24*(4), 362–370.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2), 1–27.

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *American Psychologist*, *35*(2), 151–175.

Zald, D. H., & Pardo, J. V. (1997). Emotion, olfaction, and the human amygdala: amygdala activation during aversive olfactory stimulation. *Proceedings of the National Academy of Sciences*, *94*(8), 4119–4124.

Zanna, M. P., & Cooper, J. (1974). Dissonance and the pill: an attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology*, *29*, 703–709.

Ziegler, R., & Burger, A. M. (2011). Mood and the impact of individuating information on the evaluation of ingroup and outgroup members: the role of mood-based expectancies. *Journal of Experimental Social Psychology*, *47*(5), 1000–1006.

Ziegler, R., (2014), Mood and processing effort: the mood-congruent expectancies approach. *Advances in Experimental Social Psychology*, *49*, 287–355.

CHAPTER 12

Implicit Social Cognition

ADAM HAHN AND BERTRAM GAWRONSKI

INTRODUCTION

The term "implicit social cognition" is conventionally used to refer to research in social psychology that uses a particular class of computerized measurement instruments to infer thoughts and affective reactions without directly asking participants to report on them. A central feature of these instruments is that they limit participants' ability to strategically control their responses, which distinguishes them from traditional instruments that rely on self-report (Gawronski & De Houwer, 2014). The measurement outcomes of these computerized instruments are commonly referred to as implicit measures, and the measurement outcomes of traditional self-report instruments are usually called explicit measures.

A common way to conceptualize the constructs of implicit social cognition refers to the idea of mental association, most notably evaluative and semantic associations (Greenwald et al., 2002). For example, the construct of attitude can be defined as the mental association between an attitude object and a positive or negative evaluation (e.g., association between *pizza* and *good*). Moreover, whereas the term "prejudice" refers to the mental association between a social group and a particular evaluation (e.g., association between *Muslims* and *bad*), the term "stereotype" can be defined as the mental association between a social group and a semantic attribute (e.g., association between *women* and *warm*). Similarly, the term "self-esteem" refers to the association between the self and a particular evaluation (e.g., association between *self* and *good*), and the *self-concept* refers to associations between the self and semantic attributes (e.g., association between *self* and *extraverted*). A valuable aspect of the concept of mental association is that it can be applied to a wide range of objects that are of interest to psychologists (e.g., consumer products, political parties). Although alternative frameworks have been proposed that reject the idea of mental association (see Hughes, Barnes-Holmes, & De Houwer, 2011), associative theorizing has been a driving force in research on implicit social cognition, including the development of measurement instruments and the generation of empirical predictions. The basic idea is that mental associations can be activated automatically, and this automatic activation in turn influences responses on the measurement instruments of implicit social cognition.

WHAT IS "IMPLICIT" ABOUT IMPLICIT SOCIAL COGNITION?

Although the term "implicit social cognition" was initially interpreted in a broader sense

(Greenwald & Banaji, 1995), it has become a descriptive label for social psychological research that uses the above-mentioned class of measurement instruments (Gawronski & Payne, 2010). However, why exactly the measurement outcomes of these instruments should be described as "implicit" is still a matter of debate. The two most prominent positions in this debate can be traced back to different historical roots of this particular research field (Payne & Gawronski, 2010).

The first line of research was inspired by cognitive research on automatic processes in attention and emerged from the desire to overcome the problems of social desirability in research using self-reports (Fazio, Jackson, Dunton, & Williams, 1995). Using sequential priming tasks (see the subsection titled "Sequential Priming Tasks"), this line of work was primarily concerned with the automatic activation of attitudes, showing that attitudes can influence evaluative responses even when participants do not have the intention to evaluate the attitude object. An important assumption underlying this research is that the impact of automatically activated attitudes on self-reports is reduced when participants are motivated and able to control their responses (Fazio, 2007). Thus, in this line of work, the implicit-explicit distinction is typically used to describe different kinds of measurement instruments, in that implicit measures are conceptualized as instruments that limit the opportunity for strategic control and explicit measures are conceptualized as instruments that permit strategic control.

The second line of research, also concerned with the lack of honest self-reports, grew out of cognitive research on implicit memory (Greenwald & Banaji, 1995). The central assumption underlying this work is that prior experiences can influence responses even when participants are unable to verbally report on those experiences. Based on this idea, Greenwald and Banaji (1995) defined implicit social cognitions as "introspectively unidentified (or inaccurately identified) traces of past experience that mediate responses" (p. 5). Although this definition specified past experiences as the inaccessible component, it has often been misinterpreted as indicating that the mental contents resulting from these experiences are inaccessible to introspection. Thus, in this line of research, the implicit–explicit distinction is typically used to distinguish between mental contents that are conscious versus unconscious. For example, whereas self-reports are assumed to reflect explicit attitudes (i.e., conscious attitudes), the new class of computerized instruments is assumed to capture implicit attitudes (i.e., unconscious attitudes).

To resolve the terminological confusion surrounding the implicit–explicit distinction, De Houwer, Teige-Mocigemba, Spruyt, and Moors (2009) suggested using the terms "implicit" and "explicit" to describe measurement outcomes rather than measurement instruments or psychological constructs. According to this conceptualization, a measurement outcome can be called implicit when the to-be-measured attribute (e.g., attitude, self-concept) influences the observed outcome in an automatic fashion (i.e., when the impact of the attribute on participants' responses is unintentional, unconscious, resource-independent, or uncontrollable; Bargh, 1994). Conversely, a measurement outcome should be called explicit when the to-be-measured attribute influences the observed outcome in a controlled fashion (i.e., when the impact of the attribute on participants' responses is intentional, conscious, resource-dependent, or controllable; Bargh, 1994). Different from the implicit versus explicit nature of measurement outcomes, measurement instruments may be described as direct or indirect, depending on whether they require a self-assessment of the to-be-measured attribute. According to this conceptualization, a measurement instrument is direct when it is based on participants' self-assessment of the to-be-measured attribute (e.g., when participants' racial attitudes are inferred from their self-reported liking of Blacks versus Whites). Conversely, a measurement instrument is indirect when it is not based on a self-assessment (e.g., when participants' racial attitudes are inferred from their speed in responding to positive and negative words after brief presentations of Black versus White faces) or when the to-be-measured attribute is inferred from a self-assessment of attributes other than the to-be-assessed attribute (e.g., when participants' racial attitudes are inferred from their self-reported liking of neutral objects after brief presentations of Black versus White faces).

MEASUREMENT INSTRUMENTS

Measurement instruments in implicit social cognition are based on the idea that automatic responses are influenced by whatever mental contents are activated upon encountering a given object. Thus, when a mental association is sufficiently strong, activation of one concept can automatically spread to other associated concepts (Collins & Loftus, 1975), and thereby influence responses on the task. For example, if a person has strong associations with Coca-Cola, seeing a can of Coca-Cola should activate these associations automatically, which should influence the person's responses to stimuli that are conceptually congruent or incongruent with these associations. Most measurement instruments in implicit social cognition make use of this logic in one way or another (Moors, Spruyt, & De Houwer, 2010).

Sequential Priming Tasks

The first type of measurement instruments in the area of implicit social cognition is based on the logic of sequential priming (for a review, see Wentura & Degner, 2010). In a typical sequential priming task, participants are briefly presented with a prime stimulus, which is followed by a target stimulus. Depending on the nature of the task, participants are asked to (1) classify the target as positive or negative (i.e., evaluative decision task), (2) classify the target in terms of a semantic property (i.e., semantic decision task), or (3) decide whether the target is a meaningful word or a meaningless letter string (i.e., lexical decision task). The basic idea underlying sequential priming tasks is that quick and accurate responses to the target should be facilitated when the target is conceptually congruent with the associations that were activated by the prime stimulus. In contrast, quick and accurate responses to the target should be impaired when the target is conceptually incongruent with the associations that were activated by the prime stimulus.

For example, if a person has strong positive associations with Pepsi, this person should be faster and more accurate in identifying the valence of positive words when he or she has been primed with the word "Pepsi" compared to priming trials with a neutral baseline stimulus (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Conversely, evaluative classifications of negative words should be slower and less accurate when the person has been primed with the word "Pepsi" compared to priming trials with a neutral baseline stimulus. Different from the focus on evaluative associations in sequential paradigms with evaluative decision tasks, sequential priming with semantic decision tasks is used to measure semantic associations.

For example, a person with strong gender-stereotypic associations should show better performance in identifying the gender of female pronouns after being presented with stereotypically female professions (e.g., nurse) than stereotypically male professions (e.g., doctor), and vice versa (Banaji & Hardin, 1996). Last, using a lexical decision task to assess racial stereotypes, a person may show facilitated classifications of target words related to positive and negative stereotypes of African Americans (e.g., athletic, criminal) after being primed with Black faces compared to priming trials with a neutral baseline stimulus (Wittenbrink, Judd, & Park, 1997).

Sequential priming tasks have been used with supraliminally (e.g., Fazio et al., 1995) as well as subliminally presented primes (e.g., Wittenbrink et al., 1997). However, although widely used, sequential priming tasks have been criticized for their low reliability, which rarely exceed Cronbach's alpha values of .50 (Gawronski & De Houwer, 2014). This limitation has led researchers to develop alternative instruments that show reliability estimates that are comparable to the ones of traditional self-report measures.

Implicit Association Test

The most prominent example of such measures is the Implicit Association Test (IAT), which has been developed to overcome the known limitations of sequential priming tasks (Greenwald, McGhee, & Schwartz, 1998). In the critical blocks of the IAT, participants are asked to complete two binary categorization tasks that are combined in a manner that is either congruent or incongruent with the content of the to-be-measured attribute. For example, in the commonly used race IAT, participants may be asked to categorize pictures of Black and White faces in terms of their race and positive and negative words in terms of their valence. In one critical block of the task, participants are asked to press one response key for Black faces and negative words and another response key for White faces and positive words (i.e., prejudice-congruent block). In the other critical block, participants are asked to complete the same categorization tasks with a reversed key assignment for the faces, such that they have to press one response key for White faces and negative words and the other response key for Black faces and positive words (i.e., prejudice-incongruent block). The basic idea underlying the IAT is that responses in the task should be facilitated when two mentally associated concepts are mapped onto the same response key. For example, a person who has more favorable associations with Whites than Blacks should show faster and more accurate responses when White faces share the same response key with positive words and Black faces share the same response key with negative words, compared with the reversed mapping.

IAT scores are inherently relative in the sense that they conflate four conceptually independent constructs. For example, in the race IAT, a participant's performance is jointly determined by the strength of White-positive, Black-positive, White-negative, and Black-negative associations. This limitation makes the IAT inferior to sequential priming tasks, which permit the calculation of separate priming scores if the tasks include appropriate baseline primes (see Wentura & Degner, 2010). Yet the IAT is superior in terms of its internal consistency, which is typically in the range of .70 to .90 (Gawronski & De Houwer, 2014). The latter characteristic has contributed to it being the most widely used measurement instrument in implicit social cognition.

At the same time, the IAT has been criticized for its blocked presentation of "congruent" and "incongruent" trials, which has been linked to several sources of systematic measurement error. For example, previously trained key mappings have been shown to influence performance in the second pairing in an IAT, such that IAT scores may differ depending on whether prejudice-congruent or prejudice-incongruent blocks are completed first (see Teige-Mocigemba, Klauer, & Sherman, 2010). To address these and various other limitations, researchers have developed several variants of the standard IAT that avoid blocked presentations of congruent and incongruent trials, permit nonrelative measurements for individual targets and attributes, and reduce the overall length of the task. Examples of these IAT variants include the Recoding-Free IAT (IAT-RF; Rothermund, Teige-Mocigemba, Gast, & Wentura, 2009), the Single-Block IAT (SB-IAT; Teige-Mocigemba, Klauer, & Rothermund, 2008), the Single-Category IAT (SC-IAT; Karpinski & Steinman, 2006), the Single-Attribute IAT (SA-IAT; Penke, Eichstaedt, & Asendorpf, 2006), and the Brief IAT (BIAT; Sriram & Greenwald, 2009).

Affect Misattribution Procedure

The affect misattribution procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005) was designed to combine the structural advantages of sequential priming tasks with the superior psychometric properties of the IAT (for a review, see Payne & Lundberg, 2014). Two central differences to traditional priming tasks are that (1) the target stimuli in the AMP are evaluatively ambiguous, and (2) participants are asked to report their subjective evaluations of the targets. That is, rather than inferring evaluative associations from the response time it takes a participant to decide whether a target stimulus is positive or negative, participants are presented with a neutral target stimulus and are asked to evaluate it. The basic idea is that participants may misattribute the affective feelings elicited by primes to the neutral targets and therefore judge the targets more favorably when they were primed with a positive stimulus than when they were primed with a negative stimulus. For example, in an AMP to measure racial attitudes, participants may be asked to indicate whether they find Chinese ideographs visually more pleasant or visually less pleasant than average after being primed with pictures of Black versus White faces. A preference for Whites over Blacks would be indicated by a tendency to evaluate the Chinese ideographs more favorably when the ideographs followed the presentation of a White face than when they followed the presentation of a Black face. Interestingly, priming effects in the AMP emerge even when participants are explicitly informed about the nature of the task and instructed not to let the prime stimuli influence their evaluations of the targets (Payne et al., 2005).

The AMP has been criticized for being susceptible to intentional use of the primes in evaluations of the targets (Bar-Anan & Nosek, 2012). However, the basis of this criticism has been refuted by research showing that relations between AMP effects and self-reported intentions to use the primes are due to retrospective confabulations of intentionality (i.e., participants infer that they must have had such intentions when asked afterward) rather than actual effects of intentional processes (e.g., Gawronski & Ye, 2015; Payne et al., 2013). The AMP was originally designed to measure evaluative associations, but newer versions have been developed to capture semantic associations (e.g., Sava et al., 2012).

Other Instruments

The procedures just described are the most commonly used instruments in implicit social cognition. Yet several other instruments have been developed to address specific limitations of existing tasks. We briefly describe these procedures here. (For a comprehensive review and discussions of advantages and disadvantages of different procedures, see Gawronski & De Houwer, 2014.) Many of these procedures were designed to overcome specific limitations of the IAT (e.g., relative scores, blocked structure) while preserving its psychometric advantages.

In the go/no-go association task (GNAT; Nosek & Banaji, 2001) participants are presented with different kinds of stimuli sequentially and asked to press a button ("go") in response to two types of stimuli (e.g., positive words and White faces), and to withhold a reaction ("no go") in response to all other stimuli (e.g., negative faces and non-White faces). Participants typically are given a very short response window (e.g., 600 ms), and GNAT scores are calculated in terms of accuracy (rather than response times) using signal detection theory (Green & Swets, 1966). A major advantage of the GNAT is the possibility to calculate nonrelative scores for individual target objects (e.g., attitudes toward Blacks) instead of relative scores involving two target objects (e.g., relative preference for Whites of Blacks). However, the GNAT has shown lower reliability estimates compared with the standard IAT (Gawronski & De Houwer, 2014).

On the Extrinsic Affective Simon Task (EAST; De Houwer, 2003) participants are presented with target words that are shown in two different colors (e.g., yellow and blue), as well as positive and negative words in white color. In the critical block of the task, participants are asked to respond to positive white words and words of one color (e.g., yellow) with the same key and to negative white words and words of the other color (e.g., blue) with another key (or vice versa). Because the target words are presented in different colors over the course of the task, each target is sometimes paired with the response key for positive words and sometimes with the response key for negative words. The critical question is whether participants respond faster and more accurately to the targets depending on whether they require a response with the "positive" or the "negative" key. Although the EAST eliminates the block structure of the IAT and permits a calculation of nonrelative scores for individual target objects, it has been shown to be inferior to the IAT in terms of its reliability and construct validity. This limitation has been attributed to the feature that participants do not have to process the semantic meaning of the target words (De Houwer & De Bruycker, 2007b). To address this limitation, De Houwer and De Bruycker (2007a) have developed a modified variant of the EAST that ensures semantic processing of the target words, which they called the Identification-EAST (ID-EAST).

Approach-avoidance tasks make use of the idea that positive stimuli elicit approach reactions, whereas negative stimuli elicit avoidance reactions (e.g., Brendl, Markman, & Messner, 2005; Krieglmeyer & Deutsch, 2010; Schnabel, Banse, & Asendorpf, 2006). For example, Chen and Bargh (1999) found that participants were faster at pushing a lever toward themselves (approach) in response to positive as opposed to negative stimuli. Conversely, participants were faster at pushing a leaver away from themselves (avoidance) for negative as opposed to positive stimuli (cf. Solarz, 1960). In the area of implicit social cognition, such congruency effects have been utilized to assess spontaneous responses toward a variety of objects, including social groups (e.g., Neumann, Hülsenbeck, & Seibt, 2004) and food stimuli (e.g., Seibt, Häfner, & Deutsch, 2007). In contrast to early accounts that interpreted these effects in terms of direct links between particular motor actions and motivational orientations (e.g., contraction of arm extensor = avoidance; contraction of arm flexor muscle = approach), recent research suggests that congruency effects in approach-avoidance tasks depend on the evaluative meaning that is ascribed to a particular motor action (e.g., Eder & Rothermund, 2008). Hence, responses toward the same stimuli (e.g., pulling a lever) can be reversed when the same movements are coined in negative terms (e.g., "downward") as opposed to positive terms ("pull"), and vice versa (e.g., "upward" versus "push").

In the sorting paired features task (SPFT; Bar-Anan, Nosek, & Vianello, 2009), participants are presented with pairs of stimuli (instead of just one) and provided with four instead of two response options that represent all possible combinations of stimulus types (e.g., White–good, White–bad, Black–good, and Black–bad). These response options are presented in the four corners of a computer screen and mapped onto four buttons on a computer keyboard. The specific location of the four response options is randomized over four blocks of the task. Participants' task is to quickly press the response key that captures the displayed pair of stimuli (e.g., press the key for Black-good in response to a Black face paired with a positive word). Scores are conceptualized as the difference in the response latency of accurately identifying a given combination compared to the other three combinations, standardized by each participant's individual response times across all trials. This algorithm allows for the calculation of individual rather than relative scores.

The Action Interference Paradigm (AIP; Banse, Gawronski, Rebetez, Gutt, & Morton, 2010) has been developed for research with young children for whom the demands of existing tasks might be too overwhelming. For example, using a variant of the AIP to measure gender stereotypes, Banse et al. (2010) asked children to distribute gender-stereotypical gifts (i.e., trucks and dolls) to boys and girls by pressing one of two buttons that were marked with images of a boy and a girl. In one block of the task, the children were told that the boy would like to get a truck and the girl would like to get a doll (i.e., stereotype-congruent block). In another block of the task, the children were told that the boy would like to get a doll and the girl would like to get a truck (i.e., stereotype-incongruent block). The AIP uses response latencies to measure the ease of responding similar to the IAT. Although the AIP has been developed specifically to measure gender stereotypes, procedural modifications could make it amenable for the assessment of other constructs (Gawronski & De Houwer, 2014).

Deviating from the concern with measuring associations between concepts, the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010) has been developed to measure propositional representations that capture how concepts are related. For example, in an IAT to measure self-esteem (Greenwald & Farnham, 2000), facilitated responses in the block that combines self-related words and positive words may reflect a person's actual self (i.e., *I am good*), but it may also reflect the person's ideal self (i.e., *I want to be good*). Research by Remue, Hughes, De Houwer, and De Raedt (2014) has shown that the two kinds of underlying representations are indeed conflated in the standard IAT, which can lead to theoretically implausible results (e.g., high levels of implicit self-esteem among depressed participants; see De Raedt, Schacht, Franck, & De Houwer, 2006). To overcome this limitation, the procedure includes presentations of two stimuli, such as a target object (e.g., *me*) and a valenced word (e.g., *good*). The response keys are labeled to describe different ways in which the two stimuli are related (e.g., similar versus different). Across several blocks of the task, participants are trained to learn that one key is the correct one for one type of combination and the other key is the correct one for the opposite combination. For example, participants might be trained to press the "similar" key when they are presented with the stimulus pair *I am* and *good* and the "different" key when they are presented with the stimulus pair *I am* and *bad*, or vice versa. Alternatively, participants might be trained to press the "similar" key when they are presented with the stimulus pair *I want to be* and *good* and the "different" key when they are presented with the stimulus pair *I want to be* and *bad*, or vice versa. The basic idea underlying the Implicit Relational Assessment Procedure is that responses in the task should be facilitated when a person's representation is congruent with the relation captured by the required response key than when it is incongruent with the required response. (For an alternative measure capturing propositional representations, see De Houwer, Heider, Spruyt, Roets, & Hughes, 2015).

RELATIONSHIP BETWEEN IMPLICIT AND EXPLICIT MEASURES

A common rationale for the use of implicit measures is that they provide information that cannot be captured by explicit measures. This argument is based on the observation that implicit and explicit measures tend to be weakly related. Hofmann, Gawronski, Gschwendner, Le, and Schmitt (2005) conducted a meta-analysis on the relation between IAT scores with corresponding self-reports and found an average correlation of .24. Cameron, Brown-Iannuzzi, and Payne (2012) found similar results in a meta-analysis on sequential priming tasks. However, in both cases there was also considerable variation in correlations, depending on the domain studied as well as procedural and methodological factors. Overall, correlations between implicit and explicit measures tend to be larger for self-reported judgments of feelings and affect compared to more cognitive judgments (e.g., Gawronski & LeBel, 2008; Smith & Nosek, 2011). For example, in a study by Banse, Seise, and Zerbes (2001), scores of a gay–straight IAT showed higher correlations to self-reported affective reactions toward gay people (e.g., self-reported affect when seeing two men kissing each other) compared to self-reported cognitive reactions (e.g., agreement with the statement that gay men should not be allowed to work with children). Implicit and explicit measures also show higher correlations when participants are given less time to think about their judgments than when they are encouraged to deliberate about their response (e.g., Ranganath, Smith, & Nosek, 2008). Concerning method-related factors, correlations are generally higher when implicit and explicit measures are matched in terms of their dimensionality and content. For example, implicit measures reflecting relative preferences for one group over another tend to show higher correlations to explicit measures of the same relative preference compared to explicit measures of absolute evaluations (e.g., Hofmann et al., 2005). Similarly, implicit measures reflecting evaluations of Black and White faces typically show higher correlations to explicit measures using the same faces compared to explicit evaluations of antidiscrimination policies (e.g., Payne, Burkley, & Stokes, 2008).

Different theories have been proposed to explain variations in the relation between implicit and explicit measures, two of which will be described here. Although both theories were formulated to explain relations between implicit and explicit measures in the area of attitudes, their basic assumptions can be applied to nonevaluative domains as well. (For a review, see Hofmann, Gschwendner, Nosek, & Schmitt, 2005.)

The MODE model (*M*otivation and *O*pportunity as *DE*terminants) assumes that implicit measures capture the automatic activation of attitudes in response to an object (Fazio, 2007). Depending on a person's motivation and opportunity, the person may engage in deliberate processing to scrutinize specific attributes of the object. In this case, people are assumed to base their judgments on the nature of relevant attributes instead of the automatically activated attitude. Hence, to the extent that both the motivation and the opportunity for deliberate processing are high, correlations between implicit and explicit evaluations are predicted to be low. Yet, when either the motivation or the opportunity for deliberate processing are low, people are assumed to rely on their automatic reactions, leading to higher correlations between implicit and explicit measures. These assumptions are supported by several studies indicating that evaluative judgments provided under time pressure show higher correlations with implicit measures compared to judgments provided without time pressure (e.g., Ranganath et al., 2008). Further evidence for the MODE model comes from research showing that participants with high motivation to control prejudice show lower correlations between implicit and explicit measures of racial prejudice compared to participants with a low motivation to control prejudice (e.g., Fazio et al., 1995).

Another theory that explains the relation between implicit and explicit measures is the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006, 2011). According to the APE model, implicit measures capture the behavioral outcomes of associative processes; explicit measures are assumed to reflect the behavioral outcomes of propositional processes. Associative processes are defined as the activation of mental associations on the basis of feature similarity and spatiotemporal contiguity; propositional processes are defined as the validation of the information implied by activated associations. A central assumption of the APE model is that the propositional validation of activated associations involves an assessment of consistency, in that inconsistency requires a reassessment and potential revision of one's beliefs (Gawronski, 2012). Thus, correspondence between implicit and explicit measures is assumed to depend on whether the association captured by an implicit measure is consistent with other information that is considered for a self-reported judgment. To the extent that it is consistent with other salient information, it is usually regarded as valid and therefore used as a basis for self-reported judgments. However, if it is inconsistent with other salient information, people may reject this association in order to restore cognitive consistency (Gawronski & Strack, 2004).

Although the MODE and the APE model make similar predictions in most cases, the theories differ in terms of two central assumptions. First, whereas the MODE model assumes that motivation and opportunity are the primary determinants of implicit–explicit relations, the APE model proposes cognitive consistency as the central proximal factor. To illustrate this difference, consider Fazio et al.'s (1995) finding that the relation between implicit and explicit measures of prejudice is higher for participants with low motivation to control prejudice compared to participants with high motivation to control prejudice. From the perspective of the APE model, implicit measures of prejudice capture the affective reaction that results from the associations that are activated in response to members of the target group (e.g., negative affective reaction to African Americans resulting from negative associations). This reaction may serve as the basis for a self-reported evaluative judgment (e.g., *I dislike African Americans*), unless such a judgment would be inconsistent with other salient information. In the case of racial prejudice, other salient information may include egalitarian beliefs (e.g., *Negative evaluations of disadvantaged groups are wrong*) and beliefs about discrimination (e.g., *African Americans represent a disadvantaged group*). According to APE model, consistency among these beliefs may be restored by rejecting one's affective reaction as a basis for a self-reported evaluative judgment (e.g., *I like African Americans*). Yet consistency may also be restored by changing one's egalitarian beliefs (e.g., *Negative evaluations of disadvantaged groups are okay*) or one's beliefs about discrimination (e.g., *African Americans do not represent a disadvantaged group*). These considerations lead to the novel prediction that strong egalitarian beliefs (i.e., high motivation to control prejudice) should be insufficient to reduce the relation between implicit and explicit measures of racial prejudice when participants maintain cognitive consistency by denying racial discrimination. In this case, a person may report negative feelings toward African Americans and nevertheless maintain the belief that one should not express negativity toward disadvantaged groups, because the person denies that African Americans represent a disadvantaged group (akin to the concept of "modern racism"; McConahay, 1983). This prediction has been confirmed by Gawronski, Peters, Brochu, and Strack (2008), who found high correlations between implicit and explicit measures of racial prejudice when either egalitarian beliefs or perceived discrimination were low. Correlations between the two measures were reduced only when both egalitarian beliefs and perceived discrimination were high (see also Brochu, Gawronski, & Esses, 2011). These results suggest that cognitive consistency functions as the primary proximal determinant of implicit–explicit relations, whereas motivation and opportunity are better understood as distal determinants.

Second, whereas the MODE model assumes that deliberate processing generally reduces the relation between implicit and explicit measures, the APE model assumes that such reductions should occur only when the additionally considered information is inconsistent with the association captured by the implicit measure. To the extent that deliberate processing involves a selective search for information that supports the validity of this association, deliberate processing may in fact increase rather than decrease the relation between implicit and explicit measures. This hypothesis is consistent with research showing that selective search for information that is consistent with activated associations increases the correlation between implicit and explicit measures (e.g., Galdi, Gawronski, Arcuri, & Friese, 2012; see also Peters & Gawronski, 2011b).

PREDICTION OF BEHAVIOR

A major line of research in implicit social cognition aims to improve our understanding of psychological phenomena by using implicit measures to predict meaningful psychological outcomes (e.g., interpersonal behavior, decisions, mental health). Although the practical implications of the observed effect sizes has been the subject of debate (e.g., Greenwald, Banaji, & Nosek, 2015; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013), recent meta-analyses tend to support the predictive validity of implicit measures (e.g., Cameron et al., 2012; Greenwald, Poehlman, Uhlmann, & Banaji, 2009). According to Perugini, Richetin, and Zogmeister (2010), implicit measures may contribute to the prediction of psychological outcomes over and above explicit measures in various ways, including (1) additive patterns, (2) double-dissociation patterns, (3) moderation patterns, and (4) interactive patterns.

Additive patterns involve cases in which implicit and explicit measures of the same construct jointly predict a particular outcome. Such cases tend to emerge when implicit measures are able to capture particular aspects of the outcome that are not captured by the explicit measure. For example, in a study on the prediction of consumer behavior, Maison, Greenwald, and Bruin (2004) found that adding an implicit measure of brand preferences increased the prediction of consumer choices over and above explicit measures.

Although additive patterns have been obtained in a few studies, a more common finding is a double dissociation in the prediction of different kinds of outcomes. Many dual-process models conceptualize implicit and explicit measures in terms of different underlying processes (e.g., Fazio, 2007; Gawronski & Bodenhausen, 2006; Rydell & McConnell, 2006; Strack & Deutsch, 2004). Based on this idea, implicit measures have been claimed to be superior in the prediction of spontaneous behavior, whereas explicit measures are assumed to be superior in the prediction of deliberate behavior. (For a review, see Friese, Hofmann, & Schmitt, 2008.) In line with these assumptions, nonverbal behavior in interracial interactions has shown stronger relations with implicit as compared to explicit measures, whereas verbal behavior has been shown to reveal stronger relations to explicit as compared to implicit measures (e.g., Dovidio, Kawakami, & Gaertner, 2002). Similar findings have been obtained for the self-concept of shyness (Asendorpf, Banse, & Mücke, 2002), showing that implicit measures outperformed explicit measures in the prediction of spontaneous behavior (e.g., body posture), whereas explicit measures outperformed implicit measures in the prediction of deliberate behaviors (e.g., speech duration).

Despite the available evidence for double-dissociation patterns, several studies have shown only partial or weak dissociation patterns (Perugini et al., 2010). In these cases, implicit measures predicted spontaneous behavior and explicit measures predicted deliberate behavior, but either or both measures also predicted the respective other behavior (e.g., Richetin, Perugini, Adjali, & Hurling, 2007). From the perspective of dual-process theories, these patterns might be due to the fact that many behaviors are not cleanly classifiable as either spontaneous or deliberate but might instead have both spontaneous and deliberate elements. Thus, partial-dissociation patterns might be better described as a mixture of both additive and double-dissociation patterns in the prediction of outcomes. Based on these considerations, Perugini et al. (2010) suggested that these patterns may also be called partial additive patterns (when one measure predicts both kinds of behaviors, but the other measure predicts only one) or double additive patterns (when both implicit and explicit measures predict both spontaneous and deliberate behaviors).

Drawing on the assumptions of dual-process theories (e.g., Fazio, 2007; Strack & Deutsch, 2004), several studies have investigated factors that determine whether the same outcome is predicted by either implicit or explicit measures. Such findings can be described as reflecting a moderation pattern. The central idea underlying this research is that aspects of the person or the situation can influence the degree of control over a given behavior, which should determine whether the behavior is predicted better by either explicit or implicit measures. (For a review, see Friese et al., 2008.) Consistent with this idea, implicit measures have been shown to outperform explicit measures in the prediction of candy consumption when participants' cognitive resources were depleted. In contrast, explicit measures outperformed implicit measures in the prediction of candy consumption under control conditions, where participants presumably devoted their cognitive resources to controlling their eating behavior (e.g., Hofmann, Rauch, & Gawronski, 2007). Parallel findings have been obtained for individual differences in working memory capacity (WMC), such that eating behavior was predicted better by implicit measures for participants with low WMC, whereas the same behavior was better predicted by explicit measures for participants with high WMC (e.g., Hofmann, Gschwendner, Wiers, Friese & Schmitt, 2008). Together, these findings demonstrate how both individual differences and situational factors can determine whether implicit or explicit measures are superior in the prediction of a given behavior.

Deviating from approaches in which implicit and explicit measures are seen as competitors in the prediction of behavior, several studies have investigated interactive relations between the two kinds of measures. The central assumption underlying these studies is that discrepancies between implicit and explicit measures are indicative of an unpleasant psychological state that people aim to reduce (Rydell & McConnell, 2010). In line with this assumption, Rydell, McConnell, and Mackie (2008) found that participants who had been experimentally induced to hold discrepant evaluations of a fictitious target person on implicit and explicit measures scrutinized persuasive arguments from this person more thoroughly than participants who were induced to hold convergent evaluations. In general, people who show large discrepancies on implicit and explicit measures of a particular psychological attribute (e.g., attitude, self-concept) have been shown to process discrepancy-related information more extensively than people with small discrepancies (see also Briñol, Petty, & Wheeler, 2006). Similarly, combinations of high self-esteem on explicit measures and low self-esteem on implicit measures have been shown to predict various kinds of defensive behaviors (e.g., Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003). The basic idea behind this work is that such self-esteem discrepancies reflect a conflict between spontaneous feelings and deliberate thoughts about the self, which leads to a threatening state of insecurity that people try to overcome through various kinds of defensive behaviors (e.g., narcissistic tendencies, increased in-group bias).

Despite the available evidence for each of the four patterns, their boundary conditions are still not well understood (Perugini et al., 2010). Although many of the original predictions regarding the different patterns were derived from dual-process theories, specific predictions regarding the conditions under which each of them should occur are still lacking. Thus, an important task for future research is to identify the boundary conditions of different predictive patterns and to develop theories that explain why their occurrence depends on the identified conditions.

FORMATION, CHANGE, AND CONTEXT EFFECTS

Another central question in implicit social cognition concerns the situational determinants of variations on implicit measures.

We divide our discussion of this work into three parts that address distinct theoretical questions: (1) factors that influence the formation of mental representations, (2) factors that lead to changes in existing mental representations, and (3) context effects on the activation of existing representations. We also discuss (4) the lack of process purity of implicit measures, suggesting that some variations may be due to factors that are unrelated to the constructs of interest.

Formation

Theoretically, variations on implicit measures are best understood as reflecting the formation of a new mental representation when (1) the target object is unknown to participants and (2) the acquisition of novel information about the target object causes systematic variations on implicit measures. Empirical evidence suggests that such variations can be caused by descriptive information about an object (often called propositional learning) as well as repeated pairings between a target object and other stimuli (often called associative learning; for a review, see Gawronski & Sritharan, 2010).

The simplest example for effects of descriptive information comes from studies in which participants were given positive or negative information about unknown objects, individuals, or groups (e.g., Gregg, Seibt, & Banaji, 2006). Such effects have been shown for as a little as three statements (e.g., Gawronski, Walther, & Blank, 2005). Theoretically, these findings contradict the widespread assumption that implicit measures reflect highly overlearned associations that result from long-term socialization experiences (e.g., Rudman, 2004; Wilson, Lindsey, & Schooler, 2000). Although there is evidence that developmental factors can contribute to variations on implicit measures (e.g., Baron & Banaji, 2006; Rudman, Phelan, & Heppen, 2007), research demonstrating such rapid effects on implicit measures prohibit the reverse conclusion that variations on implicit measures could be interpreted as indicators of early life experiences.

The idea of associative learning is most prominently reflected in research on evaluative conditioning (EC). In EC, repeated pairings of a neutral conditioned stimulus (CS) with a positive or negative unconditioned stimulus (US) lead to changes in the evaluation of the CS in line with the valence of the US. (For a review, see De Houwer, Thomas, & Baeyens, 2001; for a meta-analysis, see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010.) EC effects have also been demonstrated on implicit measures (e.g., Olson & Fazio, 2001). The central assumption underlying this research is that CS–US pairings create new associations in memory, which can be captured by implicit measures.

Associative processes have also been implicated in other effects that do not involve the presentation of repeated pairings. For example, investigating effects of mere ownership with implicit measures, Gawronski, Bodenhausen, and Becker (2007) found that participants showed more favorable evaluations of newly owned objects compared to objects that they did not own. According to Gawronski et al., such ownership effects are due to a process called associative self-anchoring. The central feature of this process is that a newly owned object becomes mentally associated with the self, which leads to an associative transfer of one's self-evaluation to the owned object. To the extent that most people hold positive evaluations of themselves (e.g., Bosson, Swann, & Pennebaker, 2000; Greenwald & Farnham, 2000; Koole, Dijksterhuis, & van Knippenberg, 2001), newly owned objects should elicit more favorable responses, and these responses can be captured by implicit measures. Following a similar logic, the finding that in-groups are evaluated more favorably on implicit measures compared to out-groups has been attributed to the formation of associations between the self and one's in-group, which should lead to an associative transfer of self-evaluations to one's in-group (e.g., Roth & Steffens, 2014).

Change

Theoretically, variations on implicit measures can be understood as reflecting changes in existing mental representations when (1) the target object is well-known to participants and (2) the acquisition of new information about the target object causes systematic variations on implicit measures. In the early years of implicit social cognition, implicit measures were claimed to be much more resistant to change than explicit measures (e.g., Rudman, 2004; Wilson et al., 2000). However, this assumption has been refuted by numerous studies showing changes on implicit measures in the absence of changes on explicit measures (e.g., Gawronski & LeBel, 2008; Olson & Fazio, 2006). At the same time, there have been several demonstrations of changes on explicit measures in the absence of changes on implicit measures (e.g., Gawronski & Strack, 2004; Gregg et al., 2006). Thus, a central question in this line of research is when the acquisition of new information leads to (1) changes on implicit but not explicit measures, (2) changes on explicit but not implicit measures, and (3) corresponding changes on both explicit and implicit measures.

According to the APE model (Gawronski & Bodenhausen, 2006, 2011), changes on implicit but not explicit measures should occur when (1) a given factor influences the structure and content of associations in memory and, at the same time, (2) these newly created associations are rejected as a basis for self-reported judgments because of their inconsistency with other salient information. Resonating with the idea of associative learning in EC, this pattern has been observed most commonly when (1) a well-known CS has been repeatedly paired with a positive or negative US, presumably leading to the formation of new associations, but (2) participants rely on other information that leads them to reject the newly formed associations as a basis for their evaluative judgments of the CS (e.g., Gawronski & LeBel, 2008; Gibson, 2008; Grumm, Nestler, & von Collani, 2009; Karpinski & Hilton, 2001; Olson & Fazio, 2006). However, when participants were encouraged to rely on their affective feelings toward the CS, implicit and explicit measures typically showed corresponding effects, in that both reflected the valence of the CS–US pairings (e.g., Gawronski & LeBel, 2008; Grumm et al., 2009). The latter finding is consistent with the APE model's prediction that both implicit and explicit measures should show change when (1) a given factor influences the structure and content of associations in memory and (2) these newly created associations are accepted as a valid basis for self-reported judgments.

Another prediction of the APE model (Gawronski & Bodenhausen, 2006, 2011) is that changes on explicit but not implicit measures should occur when (1) a given factor influences the perceived validity of associations in memory and, at the same time, (2) this factor does not result in the formation of new associations. According to the APE model, this case is most likely when newly acquired information leads to inconsistency within a set of salient beliefs, and the resulting inconsistency is resolved by rejecting activated associations as a basis for self-reported judgments. Consistent with these assumptions, research by Gawronski and Strack (2004) has shown that cognitive dissonance arising from induced compliance (cf. Festinger & Carlsmith, 1959) leads to changes on explicit but not implicit measures (see also Wilson et al., 2000). The same pattern has been observed in paradigms where previously acquired information is discredited as invalid, and participants are asked to mentally reverse the previously presented information. For example, Gregg et al. (2006) presented participants with positive information about a Group A and negative information about another Group B. Next, participants were told to mentally reverse this information, such that the positive information was supposed to refer to Group B and the negative information was supposed to refer to Group A. Whereas explicit measures showed a full reversal, implicit measures reflected the content of the initial information.

A critical aspect in these studies is that the discrediting information involves a simple "negation" of activated associations, which may lead to a rejection of these associations for a judgment. Yet mere rejection of a given association for overt judgments does not necessarily lead to a deactivation of this association (see Deutsch, Gawronski, & Strack, 2006). In fact, repeated negations may often have ironic effects, in that they strengthen the associative link that is supposed to be undone. For example, rejecting the proposition "Old people are bad drivers" as false may have counterintentional effects at the associative level, in that it may strengthen the associative link between *old people* and *bad drivers*. Consistent with this hypothesis, research found that repeated negations of a stereotype enhanced (rather than reduced) the stereotypical responses on implicit measures. A successful reduction occurred only when participants repeatedly affirmed a counterstereotype (e.g., Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008). Effective changes of this kind have also been obtained in studies showing that novel evaluative information that is highly diagnostic (e.g., Cone & Ferguson, 2015) or suggests a reinterpretation of earlier information (e.g., Mann & Ferguson, 2015) effectively reverses responses on both explicit and implicit measures. According to the APE model, this pattern can be observed when (1) a given factor leads to a change in perceived validity of activated information and (2) new associations are formed by the process of propositional validation.

To summarize the different patterns that can emerge as a result of interactions between associative and propositional processes, Gawronski and Bodenhausen (2006) provided a schematic overview that includes four cases:

Case 1: A direct effect on associative representations with the newly formed associations being accepted by a propositional validity assessment. This pattern is assumed to lead to corresponding changes on implicit and explicit measures, with the change on the explicit measure being fully mediated by the change on the implicit measure (e.g., Gawronski & LeBel, 2008; Whitfield & Jordan, 2009).

Case 2: A direct effect on associative representations with the newly formed associations being rejected by a propositional validity assessment. This pattern is assumed to lead to changes on implicit but not explicit measures (e.g., Gawronski & LeBel, 2008; Olson & Fazio, 2006).

Case 3: A direct effect on the process of propositional validity assessment that leads to a rejection of activated associations. This pattern is assumed to lead to changes on explicit but not implicit measures (e.g., Gawronski & Strack, 2004; Gregg et al., 2006).

Case 4: Acquisition of new propositional information that leads to the formation of new associations. This pattern is assumed to lead to corresponding changes on implicit and explicit measures, with the change on the implicit measure being fully mediated by the change on the explicit measure (e.g., Gawronski & Walther, 2008; Whitfield & Jordan, 2009).

Additionally, when a given situation involves multiple factors with different effects, the four basic cases can also occur in various combinations. For example, opposite effects on implicit and explicit measures have been observed when repeated CS–US pairings imply an evaluation that is opposite to the one implied by newly acquired propositional information. In such cases, implicit measures have been shown to reflect the evaluation implied by the CS–US pairings, whereas explicit measures reflect the valence of the newly acquired propositional information (e.g., Moran & Bar-Anan, 2013; Rydell, McConnell, Mackie, & Strain, 2006).

Context Effects

Theoretically, variations on implicit measures can be understood as reflecting contextually induced shifts on implicit measures when (1) the target object is known to participants and (2) contextually induced variations occur in the absence of new information about the target. Consistent with the idea of contextually induced shifts, a growing body of research has shown that implicit measures are highly malleable and context dependent. (For reviews, see Blair, 2002; Gawronski & Sritharan, 2010.) For example, in research using implicit measures of racial prejudice, White participants showed more positive evaluations of Black targets when the targets were presented in the context of a church than when they were presented in the context of graffiti wall (Wittenbrink, Judd, & Park, 2001). Similarly, Roefs et al. (2006) found that evaluations of high-fat foods on implicit measures were more favorable when the foods were presented in a restaurant context than when they were presented in the context of a health clinic. Similar effects have been obtained for a wide range of contextual factors, including recently encountered members of a social group (e.g., Dasgupta & Greenwald, 2001), social roles (e.g., Richeson & Ambady, 2003), salient categories (e.g., Mitchell, Nosek, & Banaji, 2003), and mood states (e.g., Gemar, Segal, Sagrati, & Kennedy, 2001).

The context dependence of implicit measures has fueled theoretical debates as to whether implicit measures reflect stable representations in memory (e.g., Fazio, 2007) or online constructions on the basis of momentarily accessible information (e.g., Schwarz, 2007). According to representational accounts, spontaneous responses captured by implicit measures depend on how a target object is categorized. To the extent that contextual cues influence the categorization of a given object, these cues may influence which category representation is activated in response to the object, which in turn influences spontaneous responses on implicit measures. In contrast, constructivist accounts propose that spontaneous responses on implicit measures depend on momentarily accessible attributes rather than on abstract category representations. Thus, to the extent that contextual cues influence the relative accessibility of certain attributes, these cues should lead to variations in a person's responses to the same object, which should be captured by implicit measures.

Although representational and constructivist accounts attribute context effects to fundamentally different processes, either one of them can explain the available evidence for context effects on implicit measures. However, their explanations may be criticized as circular, in that they can explain any context effect in a post hoc fashion without providing testable predictions about their boundary conditions. To address this concern, Gawronski, Rydell, Vervliet, and De Houwer (2010) proposed an integrative theory that combines components of both representational and constructivist accounts. A central aim of the theory is to provide a priori predictions about the contextual conditions under which implicit measures reflect (1) initially acquired information, (2) subsequently acquired information that is inconsistent with the initial information, or (3) a mixture of both. There are two core assumptions of the theory: (1) attention to contextual cues during the encoding of evaluative information determines whether this information is stored in a context-free or contextualized representation, and (2) attention to contextual cues is typically low during the encoding of initial information, but enhanced by exposure to expectancy-violating information. Together, the two assumptions imply that initial experiences tend to be stored in context-free representations, whereas expectancy-violating information is usually stored in contextualized representations.

Applied to context effects on implicit measures, Gawronski et al.'s (2010) theory predicts that implicit measures should reflect the content of expectancy-violating information only in the context in which this expectancy-violating information was learned; whereas they should reflect the content of initially acquired information in any other context. This includes both the context in which this information was originally acquired and any novel context in which the target object has not been encountered before. These predictions have been confirmed in a series of studies by Gawronski et al., which also tested several predictions about how attentional processes can moderate the hypothesized patterns of context effects (see also Gawronski, Ye, Rydell, & De Houwer, 2014; Rydell & Gawronski, 2009). In addition to providing precise predictions about the conditions under which implicit measures should be context dependent or context independent, another contribution of the theory is that it provides novel, empirically confirmed predictions about contextual conditions under which implicit measures should change in response to novel information and under which conditions they should be resistant to change. (For a review, see Gawronski & Cesario, 2013.) The central prediction is that change should be more likely when the target object is subsequently encountered in the context in which the new information was acquired. Yet change is less likely to occur when the target object is subsequently encountered in a context that is different from the one in which the new information has been acquired. These predictions have been confirmed in several independent studies and corroborated in a recent meta-analysis (Gawronski, Hu, Rydell, Vervliet, & De Houwer, 2015).

Lack of Process Purity

In the introductory section to this chapter, we noted the fundamental role of mental associations as a core concept of implicit social cognition. In line with this idea, implicit measures are often assumed to provide direct proxies for mental associations. However, in a strict sense, implicit measures reflect behavioral outcomes, and these outcomes should not be equated with their mental underpinnings (De Houwer, Gawronski, & Barnes-Holmes, 2013). Although the impact of mental associations on implicit measures is rarely disputed in the field of implicit social cognition (for an exception, see De Houwer 2014), a considerable body of research suggests that implicit measures do not provide process-pure reflections of mental associations (Teige-Mocigemba, Klauer, & Sherman, 2010). To disentangle the contributions of multiple qualitatively distinct processes to implicit measures, theorists have developed formal models that provide quantitative estimates of these processes. These models include applications of process dissociation (Payne & Bishara, 2009), multinomial modeling (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Meissner & Rothermund, 2013; Stahl & Degner, 2007), and diffusion modeling (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007).

One of the most prominent examples is Conrey et al.'s (2005) quad model, which distinguishes among four qualitatively distinct processes underlying responses on implicit measures: (1) activation of an association, (2) detection of the correct response required by the task, (3) success at overcoming associative bias, and (4) guessing. Research using the quad model has provided more fine-grained insights into the mechanisms underlying previous findings obtained with implicit measures. Whereas some effects have been shown to be genuinely related to underlying associations, others stem from nonassociative processes, such as the ability to inhibit activated associations. (For a review, see Sherman et al., 2008). For example, whereas extended training to associate racial groups with positive or negative attributes has been shown to influence associative bias (Calanchini, Gonsalkorale, Sherman, & Klauer, 2013), alcohol-related increases in implicit measures of racial bias have been linked to impaired inhibitory control (Sherman et al., 2008). Similarly, higher scores on implicit measures of racial bias among older adults have been shown to be related to decreased ability to control associations rather than stronger negative associations compared to younger adults (Gonsalkorale, Sherman, & Klauer, 2009).

QUESTIONS AND CONTROVERSIES

Much of the popularity of implicit measures can be explained by the promise that they provide insights that cannot be gained with explicit measures (e.g., when people are either unwilling or unable to provide accurate self-reports). However, although some claims have received empirical support, others have been challenged by an accumulating body of evidence. In this section, we discuss some frequent assumptions and ongoing controversies about what implicit measures do and do not tell us.

Do Implicit Measures Uncover Unconscious Representations?

As discussed earlier in this chapter, a central component in the historical origin of implicit social cognition has been Greenwald and Banaji's (1995) definition in terms of "introspectively unidentified (or inaccurately identified) traces of past experience that mediate responses" (p. 5). Although the original definition referred to unconscious sources of mental representations, it has often been interpreted in the sense that the mental representations themselves are unconscious. The latter interpretation has become so common that many authors describe the constructs captured by implicit measures as unconscious attitudes, unconscious prejudice, unconscious stereotypes, unconscious self-esteem, and unconscious self-concepts (e.g., Bosson et al., 2000; Cunningham, Nezlek, &

Banaji, 2004; Rudman, Greenwald, Mellott, & Schwartz, 1999). These constructs are contrasted with the ones captured by explicit measures, which are often described as conscious attitudes, conscious prejudice, conscious stereotypes, conscious self-esteem, and conscious self-concepts. Empirically, the claim that implicit measures uncover unconscious representations whereas explicit measures reflect conscious representations is typically based on the low correlations between implicit and explicit measures frequently observed in this line of work (see Cameron et al., 2012; Hofmann et al., 2005).

Although it is correct that correlations between implicit and explicit measures may be low if the representations captured by implicit measures are unconscious, this valid inference does not justify the reverse conclusion that low correlations indicate an effect of unconscious representations on implicit measures (Gawronski, Hofmann, & Wilbur, 2006). After all, correlations between the two kinds of measures can be low for various reasons that have nothing to do with unconsciousness, including the motivation and opportunity to engage deliberate processing (Fazio, 2007) and cognitive inconsistency of activated mental contents (Gawronski & Bodenhausen, 2006). In fact, the unconsciousness hypothesis is at odds with the findings of studies in which participants were asked to predict their measurement scores on implicit measures. Using multiple IATs capturing attitudes toward different social groups, Hahn, Judd, Hirsh, and Blair (2014) found that participants were able to predict the patterns of their IAT scores with a high level of accuracy. Importantly, predicted and actual IAT scores were highly correlated within subjects, although traditional explicit measures showed the same low correlations with IAT scores that are typically observed in this area. These findings pose a challenge to the claim that implicit measures provide a window into unconscious representations. Yet they are consistent with theories that explain dissociations between implicit and explicit measures in terms of other processes that involve a rejection of conscious representations (e.g., Fazio, 2007; Gawronski & Bodenhausen, 2006). Neither of these theories attributes the misalignment of implicit and explicit measures to lack of awareness.

An important aspect in this context is the distinction between awareness of one's own response (introspective awareness) and awareness of how one's own response compares to the responses of other people (social awareness). To obtain a high correlation between predicted and actual measurement scores in a typical between-subjects design, participants have to know not only their own response but also where their response falls in the distribution of responses revealed by the other participants (Hahn & Gawronski, 2014; Hahn et al., 2014). Thus, a more stringent way to test the conscious versus unconscious nature of the mental representations underlying implicit measures is to investigate correlations between predicted and actual measurement scores using within-subjects designs with multiple target objects. Whereas within-subjects correlations reflect the unique role of introspective awareness in predicting a person's measurement scores (e.g., how much do I like bananas compared to oranges, apples, mangoes, etc.?), the size of between-subjects correlations is additionally influenced by social awareness (e.g., how much do I like bananas compared to the other participants in the study?). Thus, in addition to the fact that low correlations between implicit and explicit measures can be attenuated by various factors related to the processing of target information (e.g., Fazio, 2007; Gawronski & Bodenhausen, 2006), the correlations obtained in traditional between-subjects designs may underestimate the actual degree of introspective awareness when participants lack social awareness. Consistent with this concern, Hahn et al. (2014) found that participants were particularly good at predicting their measurement scores for a given target in relation to other targets, indicating high introspective awareness (within-subjects correlations in the range of .50–.60). Yet their accuracy was substantially lower for the prediction of measurement scores for a given target in relation to other participants, indicating lower social awareness (between-subjects correlations in the range of .30). In sum, evidence that participants are able to predict the patterns of their responses when asked contradicts the notion that implicit measures capture consciously inaccessible contents. Yet the prerequisites and consequences of such accurate predictions are still unclear at this point and require further investigation.

Is the Difference Between Implicit and Explicit Measures Just a Matter of Social Desirability?

A common idea underlying the use of implicit measures is that they are resistant to biasing effects of social desirability. Especially in the realm of prejudice, it is often assumed that people try to adjust their responses to social norms instead of honestly reporting their thoughts and feelings about social groups (e.g., Fazio et al., 1995). Empirically, this assumption translates into two hypotheses about implicit and explicit measures (Gawronski, LeBel, & Peters, 2007). First, the correspondence between implicit and explicit measures should be moderated by social desirability, such that the correlation between implicit and explicit measures should decrease as a function of increasing social desirability concerns. Second, it should be difficult to impossible for participants to strategically influence their scores on implicit measures.

Research examining the first hypothesis has produced mixed results. On one hand, Nosek (2005) found a significant relation between self-presentational concerns and the magnitude of correlations between implicit and explicit evaluations across 57 different attitude objects. On the other hand, a meta-analysis by Hofmann et al. (2005) did not find any relation between implicit–explicit correlations and the level of social desirability associated with a given content domain. Research that used individual difference measures of socially desirable responding (e.g., Crowne & Marlow, 1960) also failed to find the predicted relation with the magnitude of implicit–explicit correlations (e.g., Egloff & Schmuckle, 2003; Hofmann, Gschwendner, & Schmitt, 2005). More supportive evidence comes from research that has investigated self-presentational concerns for particular content domains rather than domain-independent concerns with socially desirable responding. For example, in the domain of prejudice, several studies found lower correlations between implicit and explicit measures for participants with a high motivation to control prejudice than for participants with a low motivation to control prejudice (e.g., Degner & Wentura, 2008; Fazio et al. 1995; Gawronski, Geschke, & Banse, 2003; Payne et al., 2005). Yet, as we outlined earlier in this chapter, even this pattern is limited to certain conditions, in that motivation to control prejudice has been shown to reduce implicit–explicit correlations only for participants who perceive high levels of discrimination, but not for participants who perceive low levels of discrimination (e.g., Brochu et al., 2011; Gawronski et al., 2008). The latter finding suggest that, although motivational factors do influence the correspondence between implicit and explicit measures, their impact is more distal and mediated by cognitive consistency as a proximal factor (see Gawronski & Bodenhausen, 2006, 2011). Thus, although motivational factors contribute to dissociations between implicit and explicit measures, this conclusion does not permit the opposite conclusion that dissociations between implicit and explicit measures generally reflect a bias of dishonest or socially desirable responding on explicit measures.

Research examining the second hypothesis has also produced mixed results. On one hand, several studies show that instructions to bias or "fake" one's responses do not affect the scores of implicit measures (e.g., Asendorpf, Banse, & Mücke, 2002; Egloff & Schmukle, 2002; Steffens, 2004). On the other hand, research using the quad model has shown that variations in measurement scores are significantly related to differences in the success of overcoming associative biases (Conrey et al., 2005), which has been linked to a variety of individual differences and contextual factors. (For a review, see Sherman et al., 2008.) Overall, the available evidence to date suggests that, although implicit measures are less susceptible to strategic influences than explicit measures, implicit measures are not entirely immune to strategic control. Yet such influences seem to depend on a number of conditions, such as the use of particular response strategies (e.g., Teige-Mocigemba & Klauer, 2008), sufficient time (e.g., Degner, 2009), and prior experience with the task (e.g., Fiedler & Bluemke, 2005).

An important issue in this context is whether strategic influences involve either reactive control of one's responses on the task or proactive control of the mental contents that influence one's responses on the task (see Gawronski, LeBel, et al., 2007). Most research on "faking" effects has focused on reactive control of overt responses. The overall conclusion that can be drawn from this research is that reactive control is difficult but not impossible. Interestingly, research on proactive control has typically found a strong susceptibility of implicit measures to intentional influences. For example, in one of the first studies on this question, Blair, Ma, and Lenton (2001) found reduced scores on a gender-stereotyping IAT for participants who were asked to think vividly about counterstereotypical exemplars. Expanding on this finding, Peters and Gawronski (2011b) showed that recall of specific autobiographical memories can influence self-concept scores on an introversion–extraversion IAT, and this effect emerged regardless of whether participants were directly instructed to recall specific memories or the content of recalled memories was manipulated by making certain memories more desirable.

Together, the available evidence suggests that responses on implicit measures can be influenced through proactive control strategies involving the intentional activation of specific mental contents. Yet reactive control of one's responses on the task seems to be more difficult and contingent on various boundary conditions. Together, these findings show that effective control of implicit measures requires more elaborate strategies than control of explicit measures. However, they contradict the simplified notion that implicit measures are generally immune to strategic influences.

Do Implicit Measures Capture a Person's True Beliefs or Just Cultural Associations?

Another frequent question about implicit measures is whether they reflect a person's true beliefs or just culturally shared associations. The former interpretation resonates with the idea that implicit measures are less susceptible to strategic control than explicit measures (e.g., Fazio et al., 1995). The latter interpretation is based on the idea that implicit measures might be influenced by incidental aspects of one's cultural environment that are not reflected in explicit measures (e.g., Arkes & Tetlock, 2004). To evaluate the validity of the competing views, we deem it important to distinguish between a philosophical and an empirical aspect of the debate.

The philosophical aspect concerns the question of which type of behavior should be regarded as reflecting a person's true self. On one hand, there is the view that a person's true self is revealed when intentional control over one's responses fails. On the other hand, there is the equally plausible view that a person's true self is reflected in what the person consciously intends to do or say. Whereas the first interpretation equates the true self with uncontrolled behavior, the second interpretation equates the true self with intentionally controlled behavior. To the extent that implicit measures capture responses under conditions of limited control and explicit measures capture intentionally controlled responses, the two philosophical views have conflicting implications about whether either implicit or explicit measures reflect a person's true self (Gawronski, Peters, & LeBel, 2008). However, because the preference for either of the two interpretations is a matter of philosophical worldviews rather than empirical observation, any claims about the true self depend on one's subjective preference for one or the other view. Thus, even though responses on implicit measures clearly fall into the category of behavior with limited control, any depiction of implicit measures as revealing the true self are contingent on the subjectively preferred conceptualization of the true self.

The empirical aspect of the debate concerns the questions of whether implicit measures are influenced by culturally shared associations, and, if so, whether behavior is more strongly influenced by a person's endorsed beliefs or culturally shared associations. Both questions can be answered on the basis of research reviewed in this chapter. As for the first question, research on EC suggests that implicit measures are highly sensitive to incidental pairings in the environment even when explicit measures do not show any effect of the pairings (e.g., Karpinski & Hilton, 2001; Olson & Fazio, 2006). Importantly, whether the resulting associations also influence explicit measures has been shown to depend on both the consideration of other information about the target object and the consistency of this information with the newly formed associations (e.g., Gawronski & LeBel, 2008; Grumm et al., 2009). From this perspective, the apparent conflict between the two views does not map onto two distinct types of mental associations (e.g., personal associations versus cultural associations). Instead, the debate becomes obsolete, because the endorsement of mental associations depends on the processes involved in their use for making a judgment. Moreover, the reviewed research on the prediction of behavior suggests that mental associations can influence behavior even when they are rejected as a basis for judgments and decisions. Yet, as we noted in the preceding sections, their behavioral impact is moderated by various factors related to the type of behavior (e.g., Asendorpf et al., 2002), the conditions under which the behavior is performed (e.g., Hofmann et al., 2007), and individual characteristics of the person who is performing the behavior (e.g., Richetin et al., 2007). From this perspective, the presumed boundary between two kinds of associations becomes rather blurry and difficult to defend at a conceptual level (see Gawronski, Peters, & LeBel, 2008).

Do Implicit Measures Reflect Associative or Propositional Processes?

In the introduction, we outlined that implicit social cognition as a field has been shaped by the idea that many key constructs of social psychology (e.g., attitudes, prejudice, stereotypes, self-esteem, self-concept) can be conceptualized as mental associations in memory (e.g., Greenwald et al., 2002). Expanding on this idea, an influential assumption of many dual-process theories is that implicit measures capture the behavioral outcomes of associative processes, whereas explicit measures capture the behavioral outcomes of propositional processes (e.g., Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004). Associative processes involve the activation of mental associations on the basis of feature similarity and spatiotemporal contiguity; propositional processes involve the validation of the information implied by activated associations on the basis of cognitive consistency. A central difference between the two kinds of processes is that (1) associations can be activated regardless of whether they are regarded as valid or invalid, whereas propositional reasoning is inherently concerned with the perceived validity of activated information, and (2) mental propositions capture the particular relation between objects and events whereas associations reflect mere co-occurrence information (e.g., A causes versus prevents B; A likes or dislikes B).

Although dual-process interpretations of implicit and explicit measures are very common in implicit social cognition, they have been criticized by proponents of single-process theories who argue that both implicit and explicit measures are outcomes of a single propositional process (e.g., De Houwer, 2014; Kruglanski & Gigerenzer, 2011). The most elaborate single-process account has been put forward by De Houwer (2014), who argued that implicit measures reflect the automatic formation and activation of mental propositions about the relation between events. To support this argument, De Houwer cited several studies showing that implicit measures can be influenced by verbal instructions and inferences (e.g., De Houwer, 2006; Gast & De Houwer, 2012) and are sensitive to information about how stimuli are related (e.g., Gawronski et al., 2005; Zanon, De Houwer, & Gast, 2012). According to De Houwer (2014), dissociations between implicit and explicit measures occur because implicit measures involve constrained processing conditions during the retrieval of information, not because they tap into two distinct processes or representations. Whereas some information may be activated quickly without requiring a lot of cognitive effort, other information may require time and cognitive resources to be retrieved from memory. Thus, whereas the former type of information should have a strong effect on implicit measures, the latter type of information may influence only explicit, but not implicit, measures. Similar ideas have been advanced by researchers who emphasize the temporal dynamics of information activation and information integration in the course of generating an evaluative response (e.g., Cunningham, Zelazo, Packer, & Van Bavel, 2007; Wojnowicz, Ferguson, Dale, & Spivey, 2009).

In evaluating the two competing accounts, it is important to clarify the specific assumptions about which they disagree (see Gawronski, Brannon, & Bodenhausen, 2017). A central issue in this context is that effects of propositional processes on implicit measures have been addressed explicitly by dual-process theories that allow for mutual interactions between associative and propositional processes. Although it is true that some dual-process theories postulate a one-to-one mapping between processes and measures (e.g., Rydell & McConnell, 2006), other dual-process theories assume that propositional inferences can function as a distal determinant of implicit measures to the extent that they change the structure or momentary activation of associations in memory (e.g., Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004). From this perspective, effects of verbal instructions and inferences on implicit measures are consistent with dual-process accounts as long as they are in line with their assumptions about the conditions of such top-down effects. In fact, dual-process theories imply two very specific predictions about the conditions under which top-down effects of propositional inferences on associative processes should occur, allowing for direct tests between single-process and dual-process accounts.

First, dual-process theories predict that information about the validity of observed stimulus contingencies should influence only explicit measures, whereas implicit measures should reflect stimulus contingencies regardless of their perceived validity. This prediction stands in contrast to the one implied by single-process propositional theories, which suggest that both explicit and implicit measures should reflect the perceived validity of stimulus contingencies. Second, dual-process theories predict that information about the relation between two co-occurring stimuli (e.g., A causes versus prevents B; A likes versus dislikes B) should influence only explicit evaluations, whereas implicit measures should reflect the mere co-occurrence of stimuli regardless of their relation. To date, research on the first pair of competing hypotheses confirmed the predictions of dual-process propositional theories (e.g., Peters & Gawronski, 2011a), whereas research on the second pair of competing hypotheses has found empirical support for the predictions of single-process theories (e.g., Moran & Bar-Anan, 2013). Yet the obtained effects seem to depend on a number of boundary conditions that are not addressed by either of the two theories (e.g., Gawronski et al., 2005; Moran, Bar-Anan, & Nosek, 2015). Thus, despite the centrality of associative theorizing in the history of implicit social cognition, one of the most central questions to date concerns the nature of the processes and representations underlying implicit and explicit measures, and their implications for the debate between single-process versus dual-process theories. (For a review, see Sherman, Gawronski, & Trope, 2014.)

CONCLUSION

Research using implicit measures has provided valuable insights for many areas in psychology. Yet, as we noted throughout this chapter, there are still a number of unresolved questions that need to be addressed. Current models are well suited to explain different patterns in the prediction of behavior, but they lack specific predictions about the conditions under which a given pattern should occur. Similarly, discussions on whether the functional properties of implicit and explicit measures can be better explained by dual-process or single-process assumptions would benefit from research on the presumed roles of associative and propositional processes. We expect that both questions will play a central role in future research using implicit measures.

Another important, yet rarely acknowledged, issue is that different measurement instruments rely on different processes for the assessment of psychological attributes (Gawronski & De Houwer, 2014). Whereas some tasks are based on response interference mechanisms that involve a resolution of response conflicts (e.g., IAT), other tasks involve a disambiguation of ambiguous stimulus features (e.g., AMP). If responses on these tasks are treated as behaviors rather than as proxies for underlying mental constructs (De Houwer et al., 2013), a stronger focus on underlying mechanisms suggests that the prediction of behavior with implicit measures might depend on the overlap between the processes underlying responses on the task and the processes underlying the to-be-predicted behavior. For example, whereas the IAT may be a better predictor of behavior involving a resolution of response conflicts (e.g., inhibition of an impulse to pull the trigger of a gun in response to a Black man holding an object that is identified as harmless), the AMP may be a better predictor of behavior involving a disambiguation of ambiguous stimuli (e.g., tendency to misidentify an ambiguous object as a gun when it is held by a Black man). Hence, in addition to shedding light on the contribution of multiple distinct processes to overt responses on the tasks, a stronger focus on the mechanisms underlying implicit measures also may provide deeper insights for the prediction of behavior by implicit measures.

REFERENCES

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?" *Psychological Inquiry*, *15*, 257–278.

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: the case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, *7*, 136–141.

Banse, R., Gawronski, B., Rebetez, C., Gutt, H., & Morton, J. B. (2010). The development of spontaneous gender stereotyping in childhood: relations to stereotype knowledge and stereotype flexibility. *Developmental Science*, *13*, 298–306.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie*, *48*, 145–160.

Bar-Anan, Y., & Nosek, B. A. (2012). Reporting intentional rating of the primes predicts priming effects in the Affective Misattribution Procedure. *Personality and Social Psychology Bulletin*, *38*, 1194–1208.

Bar-Anan, Y., Nosek, B. A., & Vianello, M. (2009). The sorting paired features task: a measure of association strengths. *Experimental Psychology*, *56*, 329–343.

Bargh, J. A. (1994). The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of Social Cognition* (pp. 1–40). Hillsdale, NJ: Erlbaum.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *Psychological Record*, *60*, 527–542.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, *17*, 53–58.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242–261.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: the moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*, 828–841.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: the blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*, 631–643.

Brendl, C. M., Markman, A. B., & Messner, C. (2005). Indirectly measuring evaluations of several attitude objects in relation to a neutral reference point. *Journal of Experimental Social Psychology*, *41*, 346–368.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology*, *91*, 154–170.

Brochu, P. M., Gawronski, B., & Esses, V. M. (2011). The integrative prejudice framework and different forms of weight prejudice: an analysis and expansion. *Group Processes and Intergroup Relations*, *14*, 429–444.

Calanchini, J., Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, *43*, 321–325.

Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: a meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, *16*, 330–350.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*, 215–224.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Cone, J., & Ferguson, M. J. (2015). He did *what*? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology*, *108*, 37–57.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. (2005). Separating multiple processes in implicit social cognition: the Quad-Model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469–487.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349.

Cunningham, W. A., Nezlek, J. B., & Banaji, M. R. (2004). Implicit and explicit ethnocentrism: revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*, *30*, 1332–1346.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: a multilevel framework for attitudes and evaluation. *Social Cognition*, *25*, 736–760.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800–814.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85.

De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation*, *37*, 176–187.

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353.

De Houwer, J., & De Bruycker, E. (2007a). The identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 133–143.

De Houwer, J., & De Bruycker, E. (2007b). The Implicit Association Test outperforms the Extrinsic Affective Simon Task as an implicit measure of interindividual differences in attitudes. *British Journal of Social Psychology*, *46*, 401–421.

De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional-cognitive framework for attitude research. *European Review of Social Psychology*, *24*, 252–287.

De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The relational responding task: toward a new implicit measure of beliefs. *Frontiers in Psychology*, *6*, 319.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychological Bulletin*, *135*, 347–368.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: a review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*, 853–869.

Degner, J. (2009). On the (un)controllability of affective priming: strategic manipulation is feasible but can possibly be prevented. *Cognition and Emotion*, *23*, 327–354.

Degner, J., & Wentura, D. (2008). The extrinsic affective Simon task as an instrument for indirect assessment of prejudice. *European Journal of Social Psychology*, *38*, 1033–1043.

De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (2006). Self-esteem and depression revisited: implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy*, *44*, 1017–1028.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: negation as reflective operation. *Journal of Personality and Social Psychology*, *91*, 385–405.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*, 62–68.

Eder, A. B., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General*, *137*, 262.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, *83*, 1441–1455.

Egloff, B., & Schmukle, S. C. (2003). Does social desirability moderate the relationship between implicit and explicit anxiety measures? *Personality and Individual Differences*, *35*, 1697–1706.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, *25*, 603–637.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*, 203–210.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology*, *27*, 307–316.

Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behavior? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology*, *19*, 285–338.

Galdi, S., Gawronski, B., Arcuri, L., & Friese, M. (2012). Selective exposure in decided and undecided individuals: differential relations to automatic associations and conscious beliefs. *Personality and Social Psychology Bulletin*, *38*, 559–569.

Gast, A., & De Houwer, J. (2012). Evaluative conditioning without directly experienced pairings of the conditioned and the unconditioned stimuli. *Quarterly Journal of Experimental Psychology*, *65*, 1657–1674.

Gawronski, B. (2012). Back to the future of dissonance theory: cognitive consistency as a core motive. *Social Cognition*, *30*, 652–668.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59–127.

Gawronski, B., Bodenhausen, G. V., & Becker, A. (2007). I like it, because I like myself: associative self-anchoring and post-decisional change of implicit evaluations. *Journal of Experimental Social Psychology*, *43*, 221–232.

Gawronski, B., Brannon, S. M., & Bodenhausen, G. V. (2017). The associative-propositional duality in the representation, formation, and expression of attitudes. In R. Deutsch, B. Gawronski, & W. Hofmann (Eds.), *Reflective and Impulsive Determinants of Human Behavior* (pp. 103–118). New York, NY: Psychology Press.

Gawronski, B., & Cesario, J. (2013). Of mice and men: what animal research can tell us about context effects on automatic responses in humans. *Personality and Social Psychology Review*, *17*, 187–215.

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (2nd ed., pp. 283–310). New York, NY: Cambridge University Press.

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When "just say no" is not enough: affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology*, *44*, 370–377.

Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: associations influence the construal of individuating information. *European Journal of Social Psychology*, *33*, 573–589.

Gawronski, B., Hofmann, W., & Wilbur, C.J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, *15*, 485–499.

Gawronski, B., Hu, X., Rydell, R. J., Vervliet, B., & De Houwer, J. (2015). Generalization and contextualization in automatic evaluation revisited: a meta-analysis of successful and failed replications. *Journal of Experimental Psychology: General*, *144*, e50–e64.

Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: when implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology*, *44*, 1355–1361.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, *2*, 181–193.

Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. New York, NY: Guilford Press.

Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: a cognitive consistency perspective. *Personality and Social Psychology Bulletin*, *34*, 648–665.

Gawronski, B., Peters, K. R., & LeBel, E. P. (2008). What makes mental associations personal or extrapersonal? Conceptual issues in the methodological debate about implicit attitude measures. *Social and Personality Psychology Compass*, *2*, 1002–1023.

Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General*, *139*, 683–701.

Gawronski, B., & Sritharan, R. (2010). Formation, change, and contextualization of mental associations: determinants and principles of variations in implicit measures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 216–240). New York, NY: Guilford Press.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: dissonance changes explicit, but not implicit, attitudes. *Journal of Experimental Social Psychology*, *40*, 535–542.

Gawronski, B., & Walther, E. (2008). The TAR effect: when the ones who dislike become the ones who are disliked. *Personality and Social Psychology Bulletin*, *34*, 1276–1289.

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology*, *41*, 618–626.

Gawronski, B., & Ye, Y. (2015). Prevention of intention invention in the affect misattribution procedure. *Social Psychological and Personality Science*, *6*, 101–108.

Gawronski, B., Ye, Y., Rydell, R. J., & De Houwer, J. (2014). Formation, representation, and activation of contextualized attitudes. *Journal of Experimental Social Psychology*, *54*, 188–203.

Gemar, M. C., Segal, Z. V., Sagrati, S., & Kennedy, S. J. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology*, *110*, 282–289.

Gibson, B. (2008). Can evaluative conditioning change attitudes toward mature brands? New evidence from the Implicit Association Test. *Journal of Consumer Research*, *35*, 178–188.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, *45*, 410–414.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, *108*, 553–561.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022–1038.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*, 17–41.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*, 1–20.

Grumm, M., Nestler, S., & von Collani, G. (2009). Changing explicit and implicit attitudes: the case of self-esteem. *Journal of Experimental Social Psychology*, *45*, 327–335.

Hahn, A., & Gawronski, B. (2014). Do implicit evaluations reflect unconscious attitudes? *Behavioral and Brain Sciences*, *37*, 28–29.

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*, 1369.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychological Bulletin*, *136*, 390–421.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–1385.

Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit-explicit consistency? *European Review of Social Psychology*, *16*, 335–390.

Hofmann, W., Gschwendner, T., & Schmitt, M. (2005). On implicit-explicit consistency: the moderating role of individual differences in awareness and adjustment. *European Journal of Personality*, *19*, 25–49.

Hofmann, W., Gschwendner, T., Wiers, R., Friese, M., & Schmitt, M. (2008). Working memory capacity and self-regulation: towards an individual differences perspective on behavior determination by automatic versus controlled processes. *Journal of Personality and Social Psychology*, *95*, 962–977.

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior.

Journal of Experimental Social Psychology, *43*, 497–504.

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: propositional and behavioral alternatives. *Psychological Record*, *61*, 465–496.

Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology*, *85*, 969–978.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, *81*, 774–788.

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*, 16–32.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process-components of the Implicit Association Test: a diffusion model analysis. *Journal of Personality and Social Psychology*, *93*, 353–368.

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, *80*, 669–685.

Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach-avoidance behavior: the manikin task vs. two versions of the joystick task. *Cognition and Emotion*, *24*, 810–828.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberative judgments are based on common principles. *Psychological Review*, *118*, 97–109.

Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology*, *14*, 405–415.

Mann, T., & Ferguson, M. J. (2015). Can we undo our first impressions? The role of reinterpretation in reversing implicit evaluations. *Journal of Personality and Social Psychology*, *108*, 823–849.

McConahay, J. B. (1983). Modern racism and modern discrimination: the effects of race, racial attitudes, and context on simulated hiring decisions. *Personality and Social Psychology Bulletin*, *9*(4), 551–558.

Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the Implicit Association Test: the ReAL model for the IAT. *Journal of Personality and Social Psychology*, *104*, 45–69.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*, 455–469.

Moors, A., Spruyt, A., & De Houwer, H. (2010). In search for a measure that qualifies as implicit: recommendations based on a decompositional view of automaticity. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 19–37). New York, NY: Guilford Press.

Moran, T., & Bar-Anan, Y. (2013). The effect of object-valence relations on automatic evaluation. *Cognition and Emotion*, *27*, 743–752.

Moran, T., Bar-Anan, Y., & Nosek, B. A. (2015). Processing goals moderate the effect of co-occurrence on automatic evaluation. *Journal of Experimental Social Psychology*, *60*, 157–162.

Neumann, R., Hülsenbeck, K., & Seibt, B. (2004). Attitudes toward people with AIDS and avoidance behavior: automatic and reflective bases of behavior. *Journal of Experimental Social Psychology*, *40*, 543–550.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565–584.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*, 625–666.

Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: searching for attitudes and knowledge in implicit evaluation. *Cognition & Emotion*, *22*, 553–594.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*, 413–147.

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: personalizing the IAT. *Journal of Personality and Social Psychology*, *86*, 653–667.

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, *32*, 421–433.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*, 171–192.

Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, *20*, 272–314.

Payne, B. K., Brown-Iannuzzi, J., Burkley, M., Arbuckle, N. L., Cooley, E., Cameron, C. D., & Lundberg, K. B. (2013). Intention invention and the Affect Misattribution Procedure: reply to Bar-Anan and Nosek (2012). *Personality and Social Psychology Bulletin*, *39*, 375–386.

Payne, B. K., Burkley, M., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, *94*, 16–31.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293.

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going? In B. Gawronski, & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 1–15). New York, NY: Guilford Press.

Payne, B. K., & Lundberg, K. B. (2014). The Affect Misattribution procedure: ten years of evidence on reliability, validity, and mechanisms. *Social and Personality Psychology Compass*, *8*, 672–686.

Penke, L., Eichstaedt, J., & Asendorpf, J. B. (2006). Single Attribute Implicit Association Tests (SA-IAT) for the assessment of unipolar constructs: the case of sociosexuality. *Experimental Psychology*, *53*, 283–291.

Perugini, M., Richetin, J., & Zogmeister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 255–277). New York, NY: Guilford Press.

Peters, K. R., & Gawronski, B. (2011a). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin*, *37*, 557–569.

Peters, K. R., & Gawronski, B. (2011b). Mutual influences between the implicit and explicit self-concepts: the role of memory activation and motivated reasoning. *Journal of Experimental Social Psychology*, *47*, 436–442.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The meta-cognitive model (MCM) of attitudes: implications for attitude measurement, change, and strength. *Social Cognition*, *25*, 657–686.

Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: an exploration of the PAST Model. *Journal of Personality and Social Psychology*, *90*, 21–41.

Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology*, *44*, 386–396.

Remue, J., Hughes, S., De Houwer, J., & De Raedt, R. (2014). To be or want to be: disentangling the role of actual versus ideal self in implicit self-esteem. *PLoS ONE*, *9*(9), e108837.

Richeson, J. A., & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology*, *39*, 177–183.

Richetin, J., Perugini, M., Adjali, I., & Hurling, R. (2007). The moderator role of intuitive versus deliberative decision making for the predictive validity of implicit and explicit measures. *European Journal of Personality*, *21*, 529–546.

Roefs, A., Quaedacjers, L., Werrij, M. Q., Wolters, G., Havermans, R., Nederkoorn, C., . . . Jansen, A. (2006). The environment influences whether high-fat foods are associated with palatable or with unhealthy. *Behaviour Research and Therapy*, *44*, 715–736.

Roth, J., & Steffens, M. C. (2014). When I becomes we: associative self-anchoring drives implicit intergroup bias in minimal groups. *Social Psychology*, *45*, 253–264.

Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the influence of recoding in the implicit association test: the recoding-free implicit association test (IAT-RF). *Quarterly Journal of Experimental Psychology*, *62*, 84–98.

Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, *13*, 79–82.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. (1999). Measuring the automatic components of prejudice: flexibility and generality of the Implicit Association Test. *Social Cognition*, *17*, 437–465.

Rudman, L. A., Phelan, J. E., & Heppen, J. B. (2007). Developmental sources of implicit attitudes. *Personality and Social Psychology Bulletin*, *33*, 1700–1713.

Rydell, R. J., & Gawronski, B. (2009). I like you, I like you not: understanding the formation of context-dependent automatic attitudes. *Cognition and Emotion*, *23*, 1118–1152.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*, 995–1008.

Rydell, R. J., & McConnell, A. R. (2010). Consistency and inconsistency in implicit social cognition: the case of implicit and explicit measures of attitudes. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 295–311). New York, NY: Guilford Press.

Rydell, R. J., McConnell, A. R., & Mackie, D. M. (2008). Consequences of discrepant explicit and implicit attitudes: cognitive dissonance and increased information processing. *Journal of Experimental Social Psychology*, *44*, 1526–1532.

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science*, *17*, 954–958.

Rydell, R. J., McConnell, A. R., Strain, L. M., Claypool, H. M., & Hugenberg, K. (2007). Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information. *European Journal of Social Psychology*, *37*, 867–878.

Sava, F. A., Maricutoiu, L. P., Rusu, S., Macsinga, I., Virga, D., Cheng, C. M., & Payne, B. K. (2012). An inkblot for the implicit assessment of personality: the semantic misattribution procedure. *European Journal of Personality*, *26*, 613–628.

Schnabel, K., Banse, R., & Asendorpf, J. (2006). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: the Implicit Association Procedure (IAP). *Experimental Psychology*, *53*, 69–76.

Schwarz, N. (2007). Attitude construction: evaluation in context. *Social Cognition*, *25*, 638–656.

Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: how immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology*, *37*, 359–379.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. A., & Groom, C. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*, 314–335.

Sherman, J. W., Gawronski, B., & Trope, Y. (Eds.). (2014). *Dual-process Theories of the Social Mind*. New York, NY: Guilford Press.

Sherman, J. W., Klauer, K. C., & Allen, T. J. (2010). Mathematical modeling of implicit social cognition: the machine in the ghost. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement,* *Theory, and Applications* (pp. 156–175). New York, NY: Guilford Press.

Smith, C. T., & Nosek, B. A. (2011). Affective focus increases the concordance between implicit and explicit attitudes. *Social Psychology*, *42*, 300–313.

Solarz, A. K. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of Experimental Psychology*, *59*, 239–245.

Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental Psychology*, *56*, 283–294.

Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: a multinomial model of EAST performance. *Experimental Psychology*, *54*, 99–112.

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology*, *51*, 165–179.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247.

Teige, S., Schnabel, K., Banse, R., & Asendorpf, J. B. (2004). Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task. *European Journal of Personality*, *18*, 495–520.

Teige-Mocigemba, S., & Klauer, K. C. (2008). "Automatic" evaluation? Strategic effects on affective priming. *Journal of Experimental Social Psychology*, *44*, 1414–1417.

Teige-Mocigemba, S., Klauer, K. C., & Rothermund, K. (2008). Minimizing method-specific variance in the IAT: a single block IAT. *European Journal of Psychological Assessment*, *24*, 237–245.

Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A practical guide to Implicit Association Tests and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 117–139). New York, NY: Guilford Press.

Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 95–116). New York, NY: Guilford Press.

Whitfield, M., & Jordan, C. H. (2009). Mutual influences of explicit and implicit attitudes. *Journal of Experimental Social Psychology*, *45*, 748–759.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationships with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*, 815–827.

Wojnowicz, M. T., Ferguson, M. J., Dale, R., & Spivey, M. J. (2009). The self-organization of explicit attitudes. *Psychological Science*, *20*, 1428–1435.

Zanon, R., De Houwer, J., & Gast, A. (2012). Context effects in evaluative conditioning of implicit evaluations. *Learning & Motivation*, *43*, 155–165.

CHAPTER 13

Psychology and Neuroscience of Person Perception

JEFFREY A. BROOKS AND JONATHAN B. FREEMAN

INTRODUCTION

Understanding other people amounts to one of the most important and complex challenges humans regularly encounter. And yet, despite the enormously variable and noisy nature of social stimuli (e.g., faces and bodies, both in static and in dynamic configurations), humans reliably extract information about one another in a process that appears automatic and obligatory. Among the information spontaneously extracted from viewing even a static image of another's face are social categories, such as sex, race, and age; personality traits, such as trustworthiness and competence; emotion; identity; and intentions. The initial perceptual operations giving rise to such impressively efficient information processing are undoubtedly important and consequential for social interaction. However, most research on person perception carried out by social psychologists during the 20th century was not focused on people as perceptual targets. Instead of focusing on the role of external appearance and perceptual input (visual, auditory, etc.) that contributes to perceiving and understanding others, research typically focused on the internal states (intentions, beliefs, traits) that can be inferred about a person, the internal cognitive mechanisms responsible for processing such social knowledge, and the downstream consequences these inferences might have for interpersonal interaction (e.g., Brewer, 1988; Hassin, Bargh, & Uleman, 2002; Kunda & Thagard, 1996; for a review, see Gilbert, 1998).

Indeed, early research typically investigated person perception by presenting research participants with written descriptions of a person's behavior or personality characteristics or by directly presenting participants with social category labels and observing the kinds of assumptions and inferences made about a person based solely on category membership. Thus, within social psychology, early research focused on primarily postperceptual processes downstream of visually based perceptions, investigating the effects of placing someone in a social category on memory for and behaviors enacted toward that individual. Indeed, social categorization in particular has a long history in social psychology of being considered a precursor to the more consequential act of stereotyping (Allport, 1954; Bargh et al., 1996; Fiske & Taylor, 1991).

Although social psychologists amassed an extensive literature on the downstream consequences of person perception and social categorization, a great deal of research in cognitive psychology and neuroscience began to unravel the specific perceptual operations performed on social stimuli such as faces. Researchers in these fields were not necessarily interested in the social consequences of person perception, but understanding the process of perceiving complex stimuli, such as faces, provided an important tool for probing visual perception in general. This work largely focused on the perceptual mechanisms responsible for successful face recognition. Prominent theoretical models emerging from this research focused on the distinction between static facial cues and dynamic facial movements and the distinct perceptual processes these two types of cues elicit in the perceiver (Haxby, Hoffman, & Gobbini, 2000, 2002). These models were built off of early observations that visual processing of faces diverged between recognizing the identity of faces (invariant or static qualities of a face) and understanding dynamic qualities of the face, such as those occurring during speech production and transient displays of emotion (Bruce & Young, 1986).

Such research demonstrated the utility of integrating vision and face perception into person perception research. As research on the mechanisms of face processing developed, the field began to appreciate what a privileged status faces and bodies have in the perceptual system. Humans disproportionately deploy attention toward faces in particular, a tendency that emerges early in development. From birth, infants prefer looking at faces and facelike stimuli, and the ability to distinguish faces by sex and overt emotional expression begins to develop soon after (Nelson, 2001), suggesting at the very least an innate predisposition to attend to faces as motivationally (i.e., adaptively) relevant stimuli in the environment, although some work suggests that these early effects are due more to the perceptual features of faces rather than their functional relevance (e.g., Macchi, Turati, & Simion, 2004; Simion & Di Giorgio, 2015). Seminal work in face recognition also documented the face inversion effect (Yin, 1969), the phenomenon of extremely poor face recognition when faces are presented upside-down, which does not happen nearly as severely for nonface objects. Further research demonstrated that this unique effect occurs because faces are processed configurally, with recognition and understanding of face stimuli depending on successful encoding of the spatial layout of a face's individual features. Interestingly, bodies also exhibit an inversion effect to a similar degree as faces (Reed, Stone, Bozova, & Tanaka, 2003) and also have a similarly privileged role in sensory processing, with considerably more attention deployed to bodies than to objects in the environment (Downing, Bray, Rogers, & Childs, 2004). These results for faces and bodies were interpreted as clear evidence that they are an extremely special class of stimuli that convey a wealth of relevant information about people, and the human perceptual system is differentially tuned to their detection and understanding as a result.

Contemporary work by social psychologists and interdisciplinary researchers working from social neuroscience and "social vision" approaches have begun to integrate these insights by shifting person perception research to focus more on the early perceptual processes that underpin more sophisticated understanding of other people. Of importance to this chapter, this increased interest in the insights offered from studying visual perceptual processes was concurrent with the introduction of cognitive neuroscience techniques (particularly functional magnetic resonance imaging [fMRI]) into social psychology. The combination of new methods and theoretical approaches allowed researchers to integrate insights from vision science as well as cognitive neuroscience into their theories of person perception. This shift has depended largely on researchers integrating visual perceptual stimuli, such as faces and bodies, into their studies, allowing research on person perception to reflect the kinds of implicit and spontaneous perceptual inferences occurring in everyday social interaction.

Early work from this approach yielded the surprising insights that humans can be quite accurate in their ability to glean social information from brief glimpses of faces and bodies. Much of this work used research paradigms such as presenting participants with silent "thin slices" of a stranger's behavior, usually video clips as short as a few seconds long, and observing the inferences participants were able to make about people from these brief behavioral displays. Stunningly, participants were able to confidently rate the individuals in the video clips along a number of dimensions, such as personality (e.g., extraversion, warmth; Gangestad, Simpson, DiGeronimo, & Biek, 1992), trait anxiety (Harrigan, Harrigan, Sale, & Rosenthal, 1996), sexual orientation (Ambady, Hallahan, & Conner, 1999), and even racial bias (Richeson & Shelton, 2005). Moreover, multiple perceivers exhibit a high agreement in their judgments of a target's personality traits (termed "consensus at zero acquaintance"; Kenny, 1991), and such reliable and consensual inferences about other people often can accurately predict real-world outcomes. For example, early work on thin slices showed that judgments of college professors' nonverbal behavior from 30-second silent video clips accurately predicted the professors' end-of-semester evaluations by students (Ambady & Rosenthal, 1993; for a review, see Weisbuch & Ambady, 2011).

Inspired by the observation that humans are rapid, consistent, and often accurate in the impressions they draw from the external appearance of another person, much research has focused on the facial features that are most informative to perceivers. This work has benefited from preexisting theoretical movements in vision research known as ecological or Gibsonian approaches (Zebrowitz & Montepare, 2006). The Gibsonian approach to visual perception (Gibson, 1979) stresses the importance of directly perceptible bottom-up cues of objects in the environment and how those cues are inherently informative about the "affordances" of those objects—their function, capacity to be acted on, and possible benefits or dangers to the perceiver. A basic type of affordance would be the graspability of an object, information that is readily present in the object's stimulus features. The Gibsonian approach also allows for the influence of the perceiver in the form of "attunements"—individual differences in the sensitivity to particular affordances in the environment. The information in conscious perception thus reflects a direct perception of the world that is weighted by the adaptive function of objects in the environment as well as by an individual perceiver's attunements to which affordances are most important to pick up on. An ecological approach to person perception (McArthur & Baron, 1983) thus emphasizes the importance of bottom-up facial cues and the adaptive information they directly convey to the perceiver, such as the potential personality characteristics of an individual; whether an individual should be approached or avoided; and related opportunities for behavioral response by the perceiver (e.g., caregiving, mating). A great deal of work from this perspective has outlined the types of facial cues humans are most attuned to and how the social affordances gleaned from another person's face (e.g., whether someone looks young or old due to the invariant structure of their face) can bias impressions of their personality and behaviors enacted toward them; in this view, even bottom-up input can serve as a biasing factor in perceptions of other people (discussed later in the subsection titled “Facial Cues”).

Despite how basic and obligatory these processes seem to be, they are nonetheless malleable and susceptible to influence from a variety of factors that reside outside of sensory input. Vision science is replete with work investigating such “top-down” influences on perception, with a particular focus on the systematic biases and errors that can be caused from motivations or expectations harbored in the perceiver. This approach to vision can be traced back to early observations in psychology (e.g., Helmholtz, 1867) that sensory input often is impoverished, lacking, or brief but nonetheless must be rapidly understood and acted on by the perceiver. Thus, perception consists of the concurrent and interactive processing of ongoing sensory input with the memories, thoughts, emotions, and conceptual knowledge of the observer, together leading to a probabilistic interpretation of a given stimulus.

An influential movement in cognitive psychology beginning in the 1950s (called the New Look; Erdelyi, 1974) was the first research program to investigate these perceiver effects experimentally. For instance, researchers found that children of lower socioeconomic status tend to overestimate the size of coins in their hands (Bruner & Goodman, 1947) and individuals are more perceptually sensitive to faces when they are highly motivated to affiliate with others (Atkinson & Walker, 1956). Contemporary research also has examined the influence of transient motivational states on visual perception, showing that participants are more likely to interpret an ambiguous visual stimulus (such as two adjacent lines that could either be perceived as the letter *B* or the number *13*) in a certain way when they anticipate a positive outcome from that interpretation (Balcetis & Dunning, 2006), and ambiguous blends of two colors are more likely to be interpreted as one color instead of the other when that color is associated with financial gain for the participant (Voss, Rothermund, & Brandtstadter, 2008).

Indeed, such top-down influences on visual perception tend to have a pronounced impact when the target of perception is ambiguous (Bruner & Goodman, 1947; Pauker, Rule, & Ambady, 2010; Trope, 1986). Within the realm of person perception, this finding has important consequences for the social categorization of individuals of ambiguous social category membership (e.g., biracial individuals) or members of a perceptually ambiguous social category (e.g., gay individuals). Indeed, if top-down factors exert more of an influence when sensory input is ambiguous, then their importance to person perception cannot be understated, as social stimuli such as faces and bodies typically provide sensory input that is variable and ambiguous.

Recent research integrating perspectives from vision science, cognitive neuroscience, and social psychology has permitted investigation into the basic processes through which social information is extracted from social perceptual cues, such as faces, bodies, and voices. In this chapter, we describe the wealth of bottom-up cues and top-down factors that contribute to the process of perceiving another person. For descriptive purposes, we discuss bottom-up cues and top-down influences in separate sections, but we emphasize that there is a rich interplay between bottom-up and top-down processes during person perception, and the two may be difficult to tease apart in practice. After discussing the particular bottom-up perceptual cues and top-down factors that interact to produce stable perceptions and impressions of other people, we turn to a discussion of the neural mechanisms that have been observed to contribute to the person perception process and computational models that have been proposed to account for the complex interplay of environmental cues and top-down constraints in person perception. We end the chapter with a discussion of the downstream consequences in social interaction that can result from subtle components of these early perceptual processes.

BOTTOM-UP PERCEPTUAL CUES AND THE TARGET OF SOCIAL PERCEPTION

Facial Cues

An extensive body of research has documented the kinds of social information that can be gleaned from the human face and what specific facial cues and characteristics drive these perceptions. Indeed, person perception may be prone to systematic biases due to the underlying characteristics of face perception. Facial cues convey variant qualities of a person, such as their current emotional state and health, but they also can convey relatively invariant qualities of a person, such as sex, race, age, and personality characteristics. Social categories, particularly the "Big Three" (sex, race, and age; Brewer, 1988), have received a great deal of attention in the person perception literature because they can be gleaned reliably and efficiently from specific salient biological cues, and consequently they come to dominate our perceptions of other people.

Following advances in statistical face modeling, researchers became capable of experimentally manipulating the features of face stimuli in order to assess the differential role of certain cues in determining an individual's social category membership. Early work considered the differential impact of specific facial cues, such as the shape and size of the jaw, brows, and chin, on determining social category memberships in the domain of gender (Brown & Parrett, 1993) and age (Berry & McArthur, 1986). Specific facial cues also have been proposed to give rise to categorization along perceptually ambiguous dimensions, such as sexual orientation. Seminal work on "thin slices" of behavior found that sexual orientation could be classified reliably with above-chance accuracy from brief presentations of dynamic nonverbal cues (Ambady et al., 1999), and succeeding studies found that extremely brief exposure to static faces (e.g., 50 ms) is sufficient for accurate categorization of sexual orientation (Rule & Ambady, 2008a). These researchers additionally found that specific features of target faces—hair, eyes, and mouth—were able to yield, as a whole, relatively accurate categorizations of sexual orientation even when these cues were presented independently of any other facial features (Rule, Ambady, Adams, & Macrae, 2008). Interestingly, participants underestimated their own accuracy on these tasks, suggesting an implicit fluency for the diagnosticity of certain facial cues on seemingly ambiguous dimensions, such as sexual orientation.

Although less is known about the relative importance of specific facial cues to different social category dimensions, researchers agree that several broader qualities of the face contribute to social categorization. These patterns of features include the shape of the face (encompassing broad patterns of structural variation) as well as face pigmentation, alternately referred to as color and texture in the literature. Early studies focused on the role of face shape, and researchers theorized a primacy for face shape in driving evaluations and categorizations of faces (Biederman, 1987). Later work showed an important additional role for pigmentation cues in determining social categories such as age, race, and gender (Price & Humphreys, 1989; Hill, Bruce, & Akamatsu, 1995), but many researchers still emphasized the importance of shape, assuming that pigmentation cues (to the extent that they are informative at all) are integrated later in the perceptual process. However, recent work has shown that both patterns of shape and pigmentation cues are integrated into the social categorization process in a temporally dynamic fashion, with parallel processing of shape and pigmentation giving rise to coherent perceptions (although pigmentation actually appears to have primacy in the perception of gender, showing an influence on the perceptual process at earlier time points; Freeman & Ambady, 2011b).

Researchers also have considered the impact of a face's width-to-height ratio (fWHR), a useful measure of face shape and structure with particular importance for the perception of male faces. The fWHR is defined as the distance between the cheekbones divided by the distance between the upper lip and mid-brow, and the magnitude of the fWHR is driven by pubertal testosterone in men (Lefevre, Lewis, Perrett, & Penke, 2013). As predicted by an ecological approach to person perception (McArthur & Baron, 1983), perceivers seem to use the fWHR as a relatively accurate index of behavior. Larger fWHR is correlated with deceptive (Haselhuhn & Wong, 2012) and aggressive behaviors (Carré & McCormick, 2008), and perceivers readily evaluate individuals with high fWHR as less friendly (Hehman, Carpinella, Johnson, Leitner, & Freeman, 2013), less trustworthy (Stirrat & Perrett, 2010), and more aggressive (Carré, McCormick, & Mondloch, 2009).

With its strong suggestion of an individual's traits and behaviors, the fWHR is a prominent candidate for the type of facial cues that drive overgeneralization effects (Zebrowitz, Fellous, Mignault, & Andreoletti, 2003), prominent biases in trait attribution occurring when cues on the face are tied to specific social affordances bias impressions of another person. For example, older adults (who display lower perceived fWHR due to changes in the skin) are perceived as less aggressive, less physically capable, less socially competent, less socially dominant, and friendlier as a function of decreases in fWHR (Hehman, Leitner, & Freeman, 2014b). On the other end of the age spectrum, there is a well-documented, age-related overgeneralization effect such that adults with babyish features are in turn perceived to be more childlike (e.g., weak, submissive, vulnerable, submissive, and honest; Montepare & Zebrowitz, 1998). Central to researchers' interpretation of these effects is the Gibsonian assumption that perceivers are specifically attuned to social affordances that require rapid orienting and appropriate behavioral responses. In this view, age overgeneralization effects occur because humans are attuned to rapidly detect and act on age-related cues in the environment to provide necessary care to vulnerable infants. Similar overgeneralization effects have been observed in trait impressions from emotional expressions, such that individuals with permanent resemblance to certain canonical emotional expressions are perceived to have invariant personalities related to those traits. For example, individuals with faces resembling neutral or angry expressions usually are judged to be low in affiliative traits, while individuals with faces resembling happy expressions usually are judged to be high in affiliative traits (Montepare & Dobish, 2003).

The relationship between specific facial cues and social categories and trait attributions is further complicated by the fact that specific facial cues may signal more than one social category at a time, at times subjecting social categorization to systematic biases. For example, facial cues signaling sex-category membership overlap with cues signaling emotional state, resulting in consistent biases of impressions of trait dominance and affiliation in men and women (Hess, Adams, & Kleck, 2004, 2005). Additionally, race and sex-category membership intersect such that the African American race category shares overlapping phenotypic cues with the male sex category (Johnson, Freeman, & Pauker, 2012), biasing stereotypic expectations for individuals who do not satisfy this expected congruence (e.g., Black women; Johnson et al., 2012). However, the exact degree to which these effects are due to overlapping facial cues is under debate (Johnson et al., 2012). Another strong influence on intersectionality effects is overlap in stereotype content between categories on orthogonal dimensions (e.g., the Black race category and male sex category sharing stereotypes for aggression and athleticism). We discuss such influences on category intersectionality effects further in the subsection titled "Stereotypes."

As evident in overgeneralization effects, humans infer a wealth of trait information from features of the face. Although these inferences occur impressively rapidly and with surprising consensus in the population, they also may be particularly prone to error. A prominent model of personality characteristics is the Big Five factor model (Goldberg, 1990), which describes human personality characteristics in terms of the five factors of openness, conscientiousness, extraversion, agreeableness, and neuroticism. Early research examining the ability of perceivers to glean these traits from faces showed that they were rapid, consistent, and often accurate (e.g., Watson, 1989).

However, research on the extraction of specific personality characteristics (like the Big Five) suffers from the fact that several of these judgments are intercorrelated and difficult to differentiate in terms of their associated facial cues (Sutherland et al., 2015). Thus, a large portion of the work on personality judgments of faces has focused on broader impressions, such as trustworthiness, a personality trait which some research posits as a broader personality dimension that accounts for many of the intercorrelated judgments of more granular personality characteristics, such as those in the Big Five (Oosterhof & Todorov, 2008). Researchers studying impression formation and trait attribution from faces have offered many different possibilities for a parsimonious encoding of trait information along universal dimensions, such as warmth and competence (Fiske, Cuddy, & Glick, 2007), trustworthiness and dominance (Oosterhof & Todorov, 2008), and valence and dominance (Todorov, 2011).

These theoretical approaches are united in their attempt to account for the remarkable speed and consensus with which such judgments are made. Such theoretical accounts form a Gibsonian approach that assumes that the fundamental information extracted from faces is that which is adaptively relevant to the perceiver: In the case of warmth and competence, for example, the "warmth" dimension reflects whether a novel individual is antagonistic (approachability), and the "competence" dimension reflects whether an individual can cause harm to the perceiver (Fiske et al., 2007). Several insights have been made regarding the specific facial features that give rise to these social judgments. Motivated by work on overgeneralization effects, researchers in this dimensional approach have observed that an approachability dimension like "trustworthiness" is most closely related to a face's general resemblance to an emotional facial expression, while harm capability dimensions such as "competence" or "dominance" are most closely related to cues signifying strength and maturity (Oosterhof & Todorov, 2008). Additionally, recent work has shown that trustworthiness judgments in particular are driven by a face's averageness, such that average faces (in terms of proximity to the physical average of faces in the population) consistently are rated as more trustworthy (Sofer, Dotsch, Wigboldus, & Todorov, 2015).

Averageness is also a potent contributor to perceptions of facial attractiveness (Langlois & Roggman, 1990). Early work used face-morphing techniques to show that ratings of attractiveness for a composite face are consistently higher than ratings of attractiveness for any of the individual faces used to make the composite (Langlois & Roggman, 1990). Some researchers have proposed that this tendency is due to innate drives to pursue partners with a high degree of genetic diversity (Thornhill & Gangestad, 1993). Indeed, like personality traits, attractiveness judgments for novel faces show a high level of consensus across participants, even cross-culturally (Langlois et al., 2000). In addition to facial averageness, facial symmetry also contributes to judgments of a face's attractiveness (Grammer & Thornhill, 1994). However, other research shows that the overall symmetry of a face nevertheless correlates with attractiveness judgments when only half of a face is presented to participants (Scheib, Gangestad, & Thornhill, 1999), suggesting that symmetry may covary with other featural aspects of the face that confer attractiveness. Researchers also have examined the role of sexually dimorphic facial cues, although the emerging picture is complicated: Highly feminine cues consistently increase ratings of attractiveness for female faces (M. R. Cunningham, 1986), but male faces also are rated more attractive when some feminine cues are present on the face (Little, Burt, Penton-Voak, & Parrett, 2001).

Throughout this section, we have documented facial features that give rise to remarkably consensual judgments among perceivers in the population, including inferences about personality traits. Clearly this consistency in judgments shows that perceivers draw on some perceptual heuristic that they find to be reliable, but the issue of whether face-based trait inferences truly can be accurate is complicated. One component of the problem is the difficulty of defining perceiver accuracy for personality traits in the first place: whether to assess the correspondence between perceiver inferences and the target's self-reported personality traits, consensus of the target's peers, or real-world behavioral outcomes. For example, some research has assessed the ability of perceivers to accurately predict real-world outcomes, such as a corporate firm's success based on the facial appearance of chief executives (Re & Rule, 2016), but it remains unclear to what degree findings like this reflect accuracy per se. The complexities of this issue are covered in much greater detail elsewhere (e.g., Alaei & Rule, 2016).

In general, for accurate trait inferences to arise from static facial photos in any consistent fashion across contexts, these inferences would have to rely on cues that are relatively fixed in the target. Indeed, some work has shown that the most accurate and consistent impressions are drawn from skeletal cues, such as the fWHR (discussed previously), which drive impressions of dominance and physical ability. However, perception of other traits (such as trustworthiness) relies on a static face's resemblance to a more dynamic facial expression (such as an emotional expression), and thus these inferences are relatively less stable across multiple images of a target individual, leading to less opportunity for these trait inferences to be accurate (Hehman, Flake, & Freeman, 2015). That said, there are baseline, resting levels of such resemblances, and these potentially could be able to produce accurate judgments in a context-free fashion.

Relatedly, dynamic facial cues can convey considerably more information than static images, but research on person perception, social categorization, and face processing more broadly has focused primarily on static images of faces. There is a great deal of work on dynamic facial cues in the context of emotion perception, for which dynamic facial cues appear to convey more information about emotional state than static facial cues (Ambadar, Schooler, & Cohn, 2005). In addition, there is also a large body of work on eye gaze, a dynamic facial cue that can signal social motives and intent more generally (Frischen, Bayliss, & Tipper, 2007). We omit a deeper discussion of dynamic facial cues since they have not been studied nearly to the same extent in the context of extracting social information, such as category membership and personality traits, from faces. However, as discussed earlier in the chapter, there is a long history of research looking at dynamic bodily displays and nonverbal behavior (such as thin slices), and this body of work finds that even brief displays of nonverbal bodily behavior are extremely informative to perceivers and frequently give rise to consistent impressions that can predict real-world outcomes.

Bodily Cues

In everyday interaction, human faces are, of course, rarely perceived in isolation from a human body. As such, the body can provide a powerful source of visual context for face perception as well as an ample source of social knowledge about an individual in its own right. In the case of emotion perception, cues from the body even appear to dominate input from the face, and when body posture is incongruent with facial expressions, the ultimate emotion categorization often can be consistent with the body posture rather than the facial cues (Van den Stock, Righart, & de Gelder, 2007). As with faces, the perception of human bodies in both static and dynamic configurations is greatly privileged by the perceptual system, with perceptual attunement to bodies and "biological motion" perception (i.e., the perception of bodily movement) subject to similar privileged configural visual processes as faces (Reed et al., 2003), although these processes seem to emerge more slowly and later in development than those for faces (Freire, Lewis, Maurer, & Blake, 2006).

The perception of static bodily cues has been studied mostly in the context of emotion perception. Bodily cues are strongly suggestive of the emotional state of an individual and provide such a potent source of visual information about emotional states that they can disambiguate facial displays of emotion and also influence or override initial perceptions of facial emotion (Van den Stock et al., 2007; for a review, see de Gelder, 2005). However, the majority of research on the perception of the body more generally has examined biological motion, which primarily refers to naturalistic human movement, such as walking. The study of bodily movement was propelled by the psychophysicist Gunnar Johansson, who developed a novel technique for isolating displays of human movement from their visual context (Johansson, 1973). The stimuli created using this technique generally are referred to as point-light displays, and researchers have carried out a great deal of work examining the surprising amount of information that perceivers readily extract from these perceptually minimal stimuli. To create point-light displays, researchers attach reflective or infrared markers to an individual's major joints and head and record videos of the person in displays of naturalistic movement. When participants see these videos, only the points of light are visible.

Despite the highly impoverished and context-free nature of these stimuli, perceivers display a readiness and sensitivity to detect the information present in point-light displays. Early work showed that participants can readily and accurately identify the specific actions performed by individuals in point-light displays, such as swinging a hammer or knocking on a door (Johansson, 1973, 1975). In the case of walking, point-light displays of gait contain cues to the identity of the walker, and studies have shown that participants can reliably identify themselves and known others in point-light displays (Cutting & Kozlowski, 1977; Richardson & Johnston, 2005). Early work also showed that sex-category membership is categorized accurately in biological motion paradigms (Kozlowski & Cutting, 1977). Indeed, later work established that biological motion is a reliable and determinant cue to sex-category membership because of specific variations in male and female bodies that drive stable biomechanical variants in gait, such as the "center of moment" (Cutting, Proffitt, & Kozlowski, 1978). Fascinating work also has examined the interplay between body shape and bodily motion in driving perceptions of sexual orientation (Johnson, Gill, Reichman, & Tassinary, 2007). These researchers found that gender-atypical combinations of body shape and gait (e.g., a male body exhibiting the typical "sway" gait pattern of female bodies or a female body exhibiting the typical "swagger" gait pattern of male bodies) were consistently more likely to be categorized as homosexual. Other social categories have been studied far less in the context of biological motion, but some research suggests that age can be determined reliably from point-light displays (Montepare & Zebrowitz, 1993) as well.

Emerging work also has examined perceivers' ability to make attractiveness judgments based on point-light displays. Interestingly, these studies have largely converged with the literature on facial attractiveness, showing that cues to biological fitness (e.g., symmetry and "internal consistency"; Kluver, Hecht, & Troje, 2016) as well as the presence of sexually dimorphic cues (Troje, 2003) both contribute to judgments of attractiveness from point-light displays. In addition, studies have found that perceivers are able to extract variant psychological states from point-light displays, such as discrete emotions (Atkinson, Dittrich, Gemmell, & Young, 2004) and intent (such as whether a movement was natural or purposely exaggerated; Runeson & Frykholm, 1983). The study of biological motion provides an impressive example of the inherently social nature of perception, largely because of the frequent use of point-light displays, which are able to isolate the information extracted from motion itself regardless of the visual or social context. Overall, this work demonstrates how the visual system is surprisingly attuned to the social and informational content of specific cues from both faces and bodies.

Vocal Cues

The voice is an abundant (albeit experimentally underappreciated) source of social information as well as more basic information about the physical characteristics of the speaker. Voices are enormously informative in isolation (as when one speaks to a stranger on the phone for the first time), but they also can serve as multimodal context for the perception of someone's face (as when one finally meets someone previously only spoken to on the phone). Research on social categorizations and stereotyping largely has ignored vocal contributions, perhaps because of an implicit assumption that the voice is not as salient a cue as the face. However, this assumption may be inaccurate for some aspects of social perception, as some researchers suggest that vocal cues may be even more informative than facial cues in the case of emotion perception, due to their variability and ability to convey subtle distinctions in emotional state (e.g., a loud approach-oriented anger versus a quiet brooding anger; Scherer, 2003).

An extensive body of research has studied the basic information that is rapidly inferred from vocal cues in isolation, including physical attributes, such as height and weight (Van Dommelen, 1993), body size and shape (Evans, Neave, & Wakelin, 2006), age (Hughes & Rhodes, 2010), and affective state (Bestelmeyer, Rouger, DeBruine, & Belin, 2010), all of which are gleaned from the voice in a generally accurate manner. Moreover, this information is available to perceivers even when vocal cues are presented for extremely brief durations (Latinus & Belin, 2012). A recent study has shown that more socially consequential personality information also is rapidly extracted from vocal cues (McAleer, Todorov, & Belin, 2014). This study repeatedly presented participants with the word “hello” spoken by different targets and found that participants rapidly inferred traits such as trustworthiness, aggressiveness, competence, confidence, and attractiveness from these utterances. Highly consistent impressions of these traits were reached with exposure to vocal clips that were on average less than 400 ms in length, in keeping with the oft-reported consensus in personality judgments of strangers observed in the face perception literature. The researchers found that specific aspects of the acoustic input reliably covaried with perceived personality traits in a manner that depended on the gender of the speaker. For example, perceived dominance in male voices seems to depend on decreases in pitch, while for females perceived dominance increases with increases in the pitch of the voice. However, there were some commonalities, with the acoustic variable harmonic-to-noise ratio (indicating roughness) contributing to perceptions of valence in both male and female speakers (McAleer et al., 2014).

Vocal cues also carry social category information in a way that appears to depend on specific characteristics of the acoustic input. Gender is rapidly and accurately perceived from vocal cues, which are particularly distinct between males and females because of dimorphism in the body (Fitch & Giedd, 1999). Perceivers are very sensitive to diagnostic gender cues in the voice and can discriminate subtle differences in the femininity versus masculinity of a voice within gender categories (e.g., feminine versus masculine male voices), in a process that automatically activates relevant stereotypes (e.g., males with feminine vocal cues present in the voice are expected to be feminine and possess attributes stereotypically linked to femininity, such as sensitivity and kindness; Ko, Judd, & Blair, 2006). Research also has found that race is categorized accurately from isolated vocal cues (Walton & Orlikoff, 1994). African American voices tend to have larger frequency perturbation (varying pitch in the voice) and amplitude perturbation (varying loudness in the voice) as well as significantly lower harmonic-to-noise ratio than White voices. Participants were more successful at discriminating African American versus Caucasian speakers when these characteristics of the auditory signal were most distinct, suggesting a specific sensitivity to certain vocal cues in race perception (Walton & Orlikoff, 1994).

A growing body of research also has shown how vocal cues interact with facial cues during social categorization of faces. Notable effects have been observed in emotion perception, in which sad facial expressions are mistakenly perceived as happy when they are accompanied by a happy voice, even when participants are instructed to disregard the voice (de Gelder & Vroomen, 2000). The voice is also a salient multimodal cue in the categorization of a face's gender, and studies have shown that gender-congruent voices facilitate accurate detection of gender on a face (Smith, Grabowecky, & Suzuki, 2007) and incongruent voices can disrupt processing of face gender (Masuda, Tsujii, & Watanabe, 2005; for a review, see Campanella & Belin, 2007). Work in gender categorization also has shown that vocal cues can bias processing of gender-atypical faces (e.g., feminine male faces and masculine female faces). In one study (Freeman & Ambady, 2011c), researchers employed a computer mouse-tracking paradigm, measuring the trajectory of computer mouse movements as participants reached to click on a "male" or "female" category response in a gender categorization task. The stimuli were slightly gender-atypical male and female faces accompanied by voices that were either gender typical (e.g., a masculine male voice) or gender atypical (e.g., a feminine male voice). The researchers found that when faces were accompanied by a sex-atypical voice, participants' mouse movements continuously deviated toward the opposite category response (e.g., participants were continuously attracted to the "female" category response when categorizing a male face accompanied by a feminine male voice). This work suggests an important role for the voice as a continual source of information during person perception, interacting with and even biasing an evolving visual interpretation as participants develop a stable categorization of a face's gender (Freeman & Ambady, 2011c).

Vocal cues comprise an important source of social and nonsocial information about an individual. Humans are reliably attuned to subtle distinctions in the acoustic properties of vocal cues that signal such information, ranging from perceptions of the body size of the speaker to inferences about that individual's trustworthiness and competence. As such, vocal cues signal important information about perceptual targets in isolation and provide a source of multimodal context that can enhance or constrain the perception of faces. In such paradigms, participants integrate information present in the voice into their categorical judgments of face stimuli, even when asked to disregard the vocal input. These findings together indicate an important role for vocal cues in developing stable perceptions of other people, both on their own and embedded in the variety of bottom-up sensory cues discussed in this section. We now turn to a discussion of top-down factors, which can impact, guide, and bias the person perception process.

TOP-DOWN FACTORS IN THE ENVIRONMENT AND HARBORED IN THE PERCEIVER

Because of the abundance of information present in sensory input, the perceptual system is necessarily strategic. In addition to processing cues from the environment, person perception also must utilize the visual and social context in the environment and preexisting perceptual heuristics in the observer to make sense of ongoing sensory input. These additional top-down factors can be social or nonsocial in nature and can take the form of extraneous perceptual input (i.e., external cues in the environment that guide the perception of a target stimulus, such as the surrounding context influencing the perception of a face) or inputs from the perceiver (i.e., motivations and expectations that structure and potentially bias the processing of novel stimuli). Many of these influences on perception have not always been treated as such: Stereotypes, for example, were long considered to be triggered after the perception of an associated stimulus (Allport, 1954), while contemporary approaches appreciate the influence that stereotypes can have on a visual percept before it has fully stabilized and reached conscious awareness (Freeman & Ambady, 2011a). Indeed, a hallmark of top-down effects is that they co-occur with and even constrain basic visual processing, despite often consisting of high-level information, such as social factors that intuitively seem distinct from perception.

Visual Context

In the realm of person perception, "context" encompasses a broad range of visual aspects of the environment that provide a source of expectations and predictions about the social targets likely to be perceived in that environment. A basic and intuitive example of context in person perception is the visual scene in which a person is encountered, which is certainly relevant for determining their identity. For example, the process of deciding whether someone who looks like your boss is actually your boss likely will differ depending on whether the person is encountered in an office setting or in a nightclub. However, in person perception, "context" is multifaceted. The immediate visual context inherent to an individual (e.g., their clothes or the positioning of their body) can provide visual context for the perception of their face. A multimodal cue, such as a person's voice (discussed previously in the "Vocal Cues" subsection), can impact and even bias perception of their face, and one social category that a person belongs to can serve as context for the perception of that person's other group and category memberships due to overlapping physical cues (discussed earlier in the "Facial Cues" subsection) or stereotypes (discussed later in the "Stereotypes" subsection). Here we restrict our discussion to the influence of the immediate visual context and scene on perception of an individual's face.

Even cues inherent to the individual (e.g., hair and clothing) can be understood as a source of visual context, supplying a source of expectation and prediction for the perception of someone's face and identity. Hairstyle in particular has been studied in the context of social categorization, and studies have shown that racially ambiguous faces are more likely to be categorized as Black when they have a stereotypically Black hairstyle. Following this categorization, the faces subsequently are perceived to have more Afrocentric cues on the face (MacLin & Malpass, 2001). Clothing also can bias race perception by exerting a contextual cue to the social status of an individual, eliciting predictions about the person's race. In one study, researchers presented participants with faces morphed along a Black–White continuum, each of which was presented with low-status attire (e.g., a janitor uniform) or high-status attire (e.g., a business suit). In a mouse-tracking paradigm, participants categorized the faces as either White or Black while their computer mouse trajectories were recorded. The study found that low-status attire biased perceptions toward the Black category while high-status attire biased perceptions toward the White category. When race and status were stereotypically incongruent (e.g., a White face with low-status attire or a Black face with high-status attire), participants' mouse movements showed a continuous attraction to the opposite category, indicating that the social status associated with clothing exerted a top-down influence on the visual categorization of race. This effect was greater for more racially ambiguous faces (Freeman, Penner, Saperstein, Scheutz, & Ambady, 2011).

The visual scene also provides a useful source of prediction about the types of individuals likely to be encountered in a given environment, which in turn can bias the perception of individuals in certain contexts.

For example, salient cultural contexts provide a predictive framework for the potential races and ethnicities likely to be encountered in that context (such as a Shinto shrine eliciting the prediction that Japanese individuals will be encountered nearby). Behavioral work has shown that race categorization is facilitated by race-congruent contexts, such that Asian faces are categorized more rapidly and accurately in Asian scenic contexts, and that race-incongruent contexts interfere with race categorization (Freeman, Ma, Han, & Ambady, 2013). The scenic context also may play a particularly important role in emotion perception. Given the ambiguous and highly variable nature of emotional facial expressions, perceiving discrete emotions on another's face seems particularly reliant on contextual and multimodal cues (Barrett, Mesquita, & Gendron, 2011), as discussed earlier in the subsections on bodily (e.g., de Gelder, 2005) and voice (e.g., Scherer, 2003) cues. There is a great deal of behavioral evidence to suggest that identical facial expressions of emotion are perceived differently depending on the visual scene in which they are encountered (e.g., a neutral context, such as standing in front of a house, or a fearful context, such as a car crash; Righart & de Gelder, 2008). Similar effects occur when participants are just given prior knowledge about the social context emotional facial expressions were originally displayed in (Carroll & Russell, 1996). Social information immediately present in a scene also can influence emotion perception. For example, emotion perception can be influenced by the facial expressions of other individuals in a visual scene (Masuda et al., 2008), an effect that is impacted by perceivers' cultural differences in sensitivity to context (Ito, Masuda, & Hioki, 2012). These results are widely consistent with insights in the vision science community about the inherently predictive nature of perception, rendering these processes particularly prone to expectations guided by the environment (Bar, 2004; Summerfield & Egner, 2009).

Prior Knowledge and Familiarity

Most of the research considered thus far has dealt with the knowledge that can be inferred from a complete stranger based on their face and body. However, a great deal of early work on face perception focused instead on the ability to recognize the identity of previously encountered faces. Although taken for granted on a daily basis, the ability to rapidly recognize another person's identity is an incredible feat, given the large variety of social contexts and visual perspectives a person can be encountered in, from dim lighting, to a new haircut, to age-related changes in the face. Indeed, research has demonstrated that changes in viewpoint, lighting, and distance rarely cause difficulty in the recognition of familiar faces, but such variables greatly impede recognition of recently learned faces (Hancock, Bruce, & Burton, 2000). The visual features of a face can serve as a potent a priori source of expectations about a person, as we have observed, but familiarity with and prior knowledge about a person exert additional forces on person perception in the form of affective responding and spontaneous retrieval of semantic knowledge related to that person (Gobbini & Haxby, 2007). Classic early work showed that familiarity with an individual's identity sensitizes perception to categorically perceive that identity (Beale & Keil, 1995). Researchers morphed the faces of famous politicians (e.g., John F. Kennedy, Bill Clinton) to create a continuum of face stimuli from one individual to the other. When prompted to categorize the faces by identity, participants perceived face identity categorically, with an abrupt shift from one identity to the other despite the objectively gradient nature of the stimuli. Importantly, participants' self-reported familiarity with the identities in question predicted the degree to which they perceived the identities categorically.

Prior knowledge about an individual also can influence person perception, even if there is no firsthand familiarity with the target. In some cases, this influence seems to arise from the spontaneous extraction of person-specific semantic knowledge. For example, when participants are asked to judge the personality traits of faces, they are influenced by information previously paired with that face, even when that information (and its face pairing) is not explicitly recalled (Uleman, Blader, & Todorov, 2005). Such prior knowledge has an even more powerful biasing effect on trait attribution when the knowledge is affectively salient (e.g., knowledge that the target has previously engaged in disgusting behaviors; Todorov, Gobbini, Evans, & Haxby, 2007). However, the impact of prior knowledge also may occur at earlier processing stages, influencing visual processing via top-down attentional routes. Recent research has shown that the affective and social content about prior knowledge learned about a person can influence visual perception of their face, even at subliminal levels of processing (Anderson, Siegel, Bliss-Moreau, & Barrett, 2011). In this study, participants rehearsed face stimuli that were paired with negative, positive, or neutral information and subsequently completed a binocular rivalry task. In binocular rivalry tasks, perceptually dissimilar images are presented to the left and right visual fields of a participant, and one of the two images eventually dominates conscious visual perception (Blake, 2001). Although this dominance is not always long lasting, the duration of perceptual dominance and the particular stimulus that is most likely to dominate perception often are interpreted as evidence for a top-down attentional bias toward that particular percept. The researchers in this study found that faces previously paired with negative social affective information were more likely to dominate in binocular rivalry, and dominate for a longer period of time. Fascinatingly, this work suggests that prior knowledge about a person can become activated during preconscious processing of a target, in turn modulating attention and perception to enhance processing of the face. Thus, the perceptual system appears particularly sensitive to recognizing known others in the environment and biasing attention toward individuals who have previously been associated with socially and affectively salient knowledge. These characteristics of face perception in particular are consistent with the idea that the perceptual system is most attuned to information in the environment that is motivationally relevant to the perceiver.

Stereotypes

Stereotypes are comprised of conceptual knowledge about social groups—the types of traits, behaviors, and physical features members of a specific social group are expected to display (Allport, 1954). Stereotypes are a complex example of top-down influences on person perception since in many cases they are linked directly to specific bottom-up cues. Although all influences of stereotypes on perception depend in part to the expectations and social knowledge of the perceiver, they also rely on specific features of the environment (i.e., facial cues) that have become associated with such expectations (e.g., Afrocentric facial cues trigger stereotypes associated with the "Black" race category, including negative stereotypes, such as hostility, and more positive stereotypes, such as athleticism). As discussed at length in this chapter, when specific visual cues become linked to likely relevant outcomes in the environment, they also provide a source of prediction and expectation that influences visual processing. A fundamental aspect of stereotypes is that they are generalized to all members of a social group, and researchers traditionally assumed that these assumptions were triggered after placing an individual in a social category, as a cognitive strategy to guide effective and appropriate social interaction (Allport, 1954). However, contemporary work has come to appreciate the ability of stereotypes to continuously guide perception as well, before the process of social categorization is complete.

Social categorization has a powerful organizing effect on perception. For example, racially ambiguous faces that have been categorized as "Black" are subsequently perceived to have a darker skin tone (Levin & Banaji, 2006) and more Afrocentric facial cues (MacLin & Malpass, 2001). However, accumulating evidence suggests that this kind of top-down feedback from stereotypes to the visual system also can occur before social categorizations are complete. In one study, researchers presented participants with face stimuli that were morphed to be highly sex typical (e.g., masculine male face) or sex atypical (e.g., feminine male face). Participants were tasked with stereotyping these targets by choosing one of two adjectives (e.g., "caring," "aggressive") that they felt was most stereotypically associated with the target face. By recording the trajectories of participants' computer mouse movements en route to one of the two responses, researchers found that when participants were stereotyping atypical targets (e.g., male faces with some female cues), mouse movements continuously deviated toward the adjective stereotypically associated with the competing category (e.g., the stereotypically female adjective "caring"). When stereotyping an atypical target, the correct social category ("male") and the incorrect social category ("female") both become tentatively activated as potential ways to categorize a face. This study provided initial evidence that stereotypes associated with specific social categories also are activated before a stable categorization has been reached (Freeman & Ambady, 2009).

Stereotypes also can guide categorizations of perceptually ambiguous groups, as in the domain of sexual orientation. Some work has shown that participants utilize specific facial cues when tasked with categorizing sexual orientation (Rule & Ambady, 2008a; Rule et al., 2008), but these cues often reflect a top-down stereotypic heuristic. Specifically, in one set of studies, researchers found that faces were more likely to be categorized as gay or lesbian when a greater degree of gender "inversion" was present on the face—gender incongruency among multiple gendered facial cues (Freeman, Johnson, Ambady, & Rule, 2010). Thus, due to the culturally pervasive stereotype that gay men are feminine and lesbian women are masculine, the presence of feminine cues on a male face or masculine cues on a female face consistently biased categorization. This is a finding consistent with other research that found similar effects among multiple gendered bodily cues (Johnson et al., 2007) and is an example of the impact of top-down stereotypes on social categorization as well as the link between stereotypic assumptions and specific cues present in the environment.

Intersectionality effects, discussed briefly in the subsection titled "Facial Cues," have attracted increasing attention in the literature as an interesting example of stereotype feedback on visual perception and social categorization. Although multiple social categories certainly can intersect because of an overlap in diagnostic phenotypic cues (e.g., Afrocentric features also contain cues to the "male" sex category; Johnson et al., 2012), multiple social categories can become linked based on an overlap in stereotype associations. When cues on a face activate one social category and its associated knowledge structures and expectations (e.g., male cues on a face activating the concept "male" and associated expectations about an individual, such as "aggressive"), this facilitates the categorization of other social categories that incidentally share similar stereotypic expectations (e.g., the race category "Black," which shares the implicit expectation of aggression with the "male" category). Behaviorally, these stereotype overlaps can cause one category (e.g., "male") to facilitate recognition of another category along a different dimension (e.g., "Black"). They can provide a top-down source of prediction on the categorization of ambiguous stimuli, such that gender-ambiguous Black faces consistently are categorized as males. This overlap also means that certain individuals, such as Black females, experience stereotype incongruence during social categorization. Stereotype overlap between the "Black" and "male" categories means that Black female faces partially activate the "male" category, which impedes sex categorization (as with Asian male faces, which partially activate the "female" category; Johnson et al., 2012). Similar effects have been shown between intersecting race and emotion categories (e.g., "Black" and "angry"; Hugenberg & Bodenhausen, 2004) and sex and emotion categories (e.g., "female" and "joy"; Hess et al., 2000). In this case, the conceptual structure of stereotypes and their inherently predictive nature allow social categories and the facial cues associated with them to provide a visual context for the perception of other social categories, even those from orthogonal dimensions (for a review, see Johnson & Freeman, 2010). Indeed, a recent study showed that individual differences in stereotype overlap (e.g., similarity in conceptual content of stereotypes for the "Black" and "male" social categories) predicted the amount of perceptual intersectionality effects as measured with computer mouse-tracking (Stolier & Freeman, 2016).

Motivation

Motivation has been explored extensively as an influence on perception harbored in the perceiver. Transient motivational states can influence visual perceptions, such that individuals will perceive ambiguous stimuli in line with whatever interpretation will have a positive outcome for them (Balcetis & Dunning, 2006; Voss et al., 2008). However, motivation also can be thought of as a chronic tuning of perception toward whatever aspects of the environment are most adaptively relevant or useful, as we addressed in the introduction in our discussion of the environmental cues preferentially attended to by perceivers. Typically, these motivational influences bear weight on perception even when they reside outside of conscious awareness. For example, participants are more likely to identify an impoverished image of a gun as a gun when primed with a Black face, because of the association of African Americans with crime in the United States and the motivation to recognize and respond to potential threats in the environment (Eberhardt, Goff, Purdie, & Davies, 2004). In the realm of social categorization, researchers have shown that unconscious biological factors, such as fertility in women (Johnston, Arden, Macrae, & Grace, 2003), and conscious motivational states, such as sexual desire (Brinsmead-Stockham, Johnston, Miles, & Macrae, 2008), can modulate the process of sex categorization to increase efficiency (speed and accuracy) of the recognition of potential mates.

Individual differences in overall vigilance to certain social cues also modulate person perception. These can be thought of as individual differences in “attunements” to particular aspects of the environment, in the Gibsonian sense. For example, adults high in attachment anxiety have a speed-accuracy trade-off in perceiving emotional expressions on faces: Although they are hypervigilant to the presence of emotion cues on a face and fast to identify an emotion when a neutral face dynamically changes its expression, these assessments typically are less accurate than those made by nonanxious individuals (Fraley, Niedenthal, Marks, Brumbaugh, & Vicary, 2006). As with visual context, perceivers rely more on motivation and other perceiver inputs when the target of perception is ambiguous. (For a discussion, see Pauker et al., 2010.) For example, individuals high in stigma consciousness (the expectation of being stereotyped and stigmatized by others; Pinel, 1999) are more likely to interpret an ambiguous emotional expression as expressing contempt (Inzlicht, Kaiser, & Major, 2008). On the other end of the spectrum, White individuals high in racial prejudice are significantly more likely to categorize racially ambiguous angry faces as Black (Hugenberg & Bodenhausen, 2004; Hutchings & Haddock, 2008). Although the intersection between the emotion concept “anger” and the race category Black is common in the United States, as previously discussed, White individuals high in racial prejudice have an exaggerated version of this effect due to their particular motivation to recognize what they view as a threatening social group.

Intergroup Processes

Humans possess an intrinsic drive to form social groups and behave in ways that sustain those groups. From this tendency to support and protect the in-group while remaining suspicious of the out-group, a number of perceptual consequences emerge when categorizing an individual as an in-group or out-group member. Indeed, a growing body of research suggests a fundamental divergence in the early visual processing of in-group versus out-group faces. A substantial amount of early evidence for an in-group/out-group processing distinction came from observations of the “cross-race effect,” a robust phenomenon where recognition memory is better for own-race rather than other-race faces (Meissner & Brigham, 2001). A great deal of research looked into the underlying mechanisms of this effect, with evidence accumulating to suggest possible contributions of perceptual expertise and greater familiarity with the racial in-group (MacLin & Malpass, 2001), increased individuation (versus mere categorization) of own-race faces (Hugenberg, Young, Bernstein, & Sacco, 2010), and, importantly, poorer encoding of other-race faces due to a divergence in the way other-race faces are processed visually. In particular, findings suggest that perception of racial out-group faces relies on featural processing (encoding of isolated focal cues on a face) versus configural processing (better encoding of the gestalt spatial layout of a face; Michel, Rossion, Han, Chung, & Caldara, 2006) and that the effect may manifest at a very low level in differences in how attention is allocated to different parts of the face (Hills & Lewis, 2006, 2011; Hills & Pake, 2013). Early evidence began to suggest that this shift in processing was the fundamental mechanism behind the cross-race effect, over and above a lack of familiarity or expertise with bottom-up cues more prevalent in racial out groups. For example, researchers showed that identical racially ambiguous faces, which contain both own-race and other-race cues, were processed more holistically whenever categorized as own race (Michel, Corneille, & Roisson, 2007).

Meanwhile, research also accumulated to suggest that the cross-race effect might be one instance of a broader cross-category effect, where there is simply better encoding of in-group versus out-group faces in all cases. A powerful paradigm used to investigate these and other effects in intergroup relations is the minimal group paradigm, in which participants are assigned to an arbitrary "minimal" group with which they have no prior knowledge, familiarity, or stereotypes (Tajfel, 1970). Studies have shown that when such arbitrary coalitions are created in an experimental setting, participants rapidly adapt to this new context and even come to spontaneously categorize individuals along new arbitrary category dimensions more readily than they do by race. Researchers have speculated that these effects show that race dominates perception only because it is such a visually salient category dimension, not necessarily because it is the most important or primary of human coalitional divisions (Kurzban, Tooby, & Cosmides, 2001). Other work has shown that these effects extend to evaluative dimensions, such that participants show implicit favoritism for minimal in-group members (Van Bavel & Cunningham, 2009). Importantly, cross-category effects also have been shown for minimal groups, such that participants have better recognition memory for in-group versus out-group faces even when these groups are minimal (Van Bavel, Packer, & Cunningham, 2012). In both of these studies, the minimal groups included both Black and White individuals.

The impressive effects observed in minimal group paradigms hint at a social-cognitive influence on face perception that is driven mainly by group membership. Indeed, the salience of group identity also can impact the perception of natural groups, such as race. Chiao, Heck, Nakayama, and Ambady (2006) explored this by priming biracial individuals with either a Black or White identity. When primed with their Black identities, participants subsequently performed a visual search task with Black and White faces much like Black participants (i.e., faster visual search times for Black faces). The opposite effect was observed when participants were primed with their White identities, suggesting that top-down effects of group membership on perception can be modulated by individual differences in the salience of group identity. These effects also are influenced by enduring characteristics of the perceiver, such as individual differences in lay beliefs about traits and categories (i.e., essentialism versus incrementalism). For example, biracial individuals tend to have more flexible conceptions of group membership, which leads to differences in their reliance on context cues to categorize faces along the race dimension. Specifically, individuals belonging to only one race tend to have more essentialist conceptions of group identity, leading to a greater reliance on contextual cues (e.g., verbal race labels) to disambiguate processing of racially ambiguous faces (Pauker & Ambady 2009).

However, there also is some evidence for the role of familiarity and personal experience with out-group members in cross-category perception. An extensive literature has documented these effects by examining perceptual boundaries between social categories—the point in a continuum of morphed faces where participants perceive a transition from one category to another. Importantly, these boundaries are malleable and can be influenced by very recent context. For example, repeated exposure to male faces moves the gender category boundary toward the male category (making participants more conservative about categorizing faces as male, ultimately categorizing more faces as female; Webster, Kaping, Mizokami, & Duhamel, 2004). Similarly, the category boundary between White and Asian faces shifts toward the White category when target faces are presented in a sequence of White faces. These shifts also can be induced by more chronic changes in exposure and frequency of encounters with out-group faces. For example, Asian college students show a shift in their perceptual boundary between the White and Asian categories when they live in the United States for approximately 1 year (Webster et al., 2004). In general, researchers find that people tend to shift boundaries toward their own social identity, which potentially reflects an adaptive mechanism—humans are conservative about categorizing novel individuals as in-group members, particularly when such individuals are ambiguous in-group members, since mistakes in this categorization could be costly (Tajfel & Turner, 1979).

Finally, it is important to note the difficulty in separating genuine top-down effects on perception from top-down influences on postperceptual processes, such as response biases or recognition behavior. This issue has led some researchers to criticize the idea that top-down factors are able to penetrate perceptual representations (Firestone & Scholl, 2016), drawing on older ideas of functional modularity (Fodor, 1983; Pylyshyn, 1999). In this ongoing debate, we believe more implicit behavioral measures (such as mouse-tracking) and neuroimaging methods that can better assess perceptual representations (e.g., multivariate fMRI) and do not require explicit responses may be helpful in addressing to what extent higher-order social cognition can shape lower-level perceptual representations.

COMPUTATIONAL AND NEURAL MECHANISMS

A great deal of work in social neuroscience has aimed to determine which aspects of the brain's visual processing regions are selective for faces and bodies. This work built on prominent early models of face processing (Bruce & Young, 1986), which focused on isolating the neural systems responsible for processing invariant qualities of a face, such as a face's identity, versus variant qualities, such as dynamic facial expressions. Research on the neural systems of face perception in turn built on the incredible progress in neuroscience describing the structure of basic visual processing. After visual sensory information hits the retina, it is relayed via optic nerve fibers that largely terminate in the lateral geniculate nucleus of the thalamus. The lateral geniculate nucleus in turn relays information to primary visual cortex (striate cortex), where it initially retains its retinotopic coding (i.e., purely veridical neural processing of the light on the retina) (Bullier, 2002). Visual information is further processed by two extrastriate pathways, known as the dorsal and the ventral visual streams (Goodale & Milner, 1992). The ventral visual stream, also known as the "what" stream, comprises ventral aspects of occipitotemporal cortex, key processing regions for the visual recognition and categorization of objects. A great deal of work in social and cognitive neuroscience has focused on delineating which aspects of the ventral visual stream are selective for faces and bodies and how these neural systems interact with larger-scale brain networks. Indeed, the increasing capabilities of social and cognitive neuroscience to characterize individual cognitive functions in terms of the collaborative activity of multiple large-scale brain networks stand to revolutionize theory development in social and cognitive psychology, including the field of person perception (Barrett & Satpute, 2013).

In person perception, the ventral visual stream is primarily responsible for processing static facial cues (Haxby et al., 2000). In particular, early feature-based processing is undertaken by the occipital face area (OFA), with higher-level representation of the configural properties of a face occurring in the fusiform gyrus (FG). The FG in particular has been the focus of intense scrutiny regarding its precise role in face processing. This debate follows from the discovery of a functionally defined region in the FG commonly referred to as the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997), which appears strongly selective for faces and has been posited as a face perception module. Regardless of the specific computations performed by the FFA and its status as a truly distinct functional module, the FG/FFA has undisputed primacy in the neural processing of faces (Haxby et al., 2000). Moreover, because of the ventral visual stream's role in categorizing visual objects in general, and its particular sensitivity to faces, it is unsurprising that these regions are routinely implicated in social category representation. Multivoxel pattern analyses (MVPA) of fMRI data, which are able to isolate the unique patterns of neural activity associated with specific stimulus conditions (Norman, Polyn, Detre, & Haxby, 2006), are consistently able to isolate specific representations for social categories in the race (Contreras, Banaji, & Mitchell, 2013) and sex (Kaul, Rees, & Ishai, 2011) dimensions. Similarly, processing of static bodily cues is subserved by the extrastriate body area and fusiform body area (Peelen & Downing, 2007). However, when bodies provide a disambiguating context for the social categorization of faces, the meaningful social category information still appears to be encoded in the FG/FFA (Cox, Meyers, & Sinha, 2004). Additionally, fascinating work shows that gender-specific olfactory cues (i.e., compounds that mimic sex hormones) also elicit FG activity (Savic, Berglund, Gulyas, & Roland, 2001), suggesting a more general sensitivity to social category information in the FG.

The ventral visual stream is also sensitive to several of the top-down factors discussed previously. An important aspect of visual processing is that sensory input becomes more constrained by conceptual and predictive factors along more anterior aspects of the ventral visual stream (Grill-Spector & Weiner, 2014). Although early feature processing occurs in the OFA, more anterior regions, such as the FG, involved in forming higher-order perceptual characteristics of a face, largely are constrained by dense structural and functional connections with higher-order regions, such as the orbitofrontal cortex (OFC), which comprises the most ventral aspects of the prefrontal cortex (PFC; Zanto, Rubens, Thangavel, & Gazzaley, 2011). Some evidence also suggests that the anterior temporal lobe, at the end of the ventral visual processing stream, houses amodal representations of face identity (Anzellotti & Caramazza, 2014). However, even the FG appears to be sensitive to complex information about a face's identity (e.g., prior knowledge and familiarity; Rotshtein, Henson, Treves, Driver, & Dolan, 2005). Moreover, the FG shows an in-group/out-group distinction even when the groups in question are minimal groups, suggesting that intergroup effects in the FG reflect higher-level coding of perceptual targets beyond visual features or familiarity effects (Van Bavel et al., 2008). It is worth noting that in this case, "in-group/out-group distinction" simply means that there was greater activity in the right FG. Since FG activity reflects higher-level processing of faces, including configural processing and enhanced encoding, this simple distinction in activity could suggest a potential contributing mechanism to cross-category effects. However, examining neural activity just in terms of overall mean activation across a brain region sometimes obscures a more nuanced story. A recent fMRI study divided participants into mixed-race minimal groups and showed that even though the FG shows increased activity for minimal in-group faces, face race still is representationally distinct in the FG as measured by MVPA (Ratner, Kaul, & Van Bavel, 2013). MVPA also reveals an interesting pattern of effects for individuals who are implicitly prejudiced, who show more distinct (i.e., dissimilar) neural representations of own- and other-race faces in the FG (Brosch, Bar-David, & Phelps, 2013). Since evaluative bias causes individuals to see out-group members as more similar (and thus more different from in-group members; Ostrom & Sedikides, 1992), these findings suggest that these individual differences in evaluative bias are detectable at the level of neural representations of social categories and reflect a perceptual bias.

Recent work also has used MVPA approaches to examine how stereotype intersectionality effects can bias neural representations of faces in the FG and OFC. Importantly, the OFC is theorized to constrain face representations in the FG via automatic activation of stereotypes and associated predictions and expectations (Knutson, Mah, Manly, & Grafman, 2007; Milne & Grafman, 2001). In a recent set of studies, researchers presented participants with faces crossed on gender, race, and emotion categories. Participants also completed a mouse-tracking task that indexed individual differences in perceptual biases (e.g., subjective perceptions of similarity between the "Black" and "male" categories) as well as a stereotype contents task to assess individual differences in conceptual knowledge (i.e., stereotypes) about each social category. Categories exhibiting greater conceptual similarity (e.g., Black and Angry) predicted greater perceptual biases, and this biased similarity was in turn reflected in the representational similarity of categories' patterns of neural activity in the FG and OFC. For example, a participant whose stereotypes overlapped between the "Black" and "male" categories tended to be see faces belonging to those categories more similarity (perceptual biases assessed with mouse-tracking), and FG and OFC neural patterns representing the "Black" and "male" categories were consistently more similar, even when controlling for any bottom-up physical similarity (Stolier & Freeman, 2016). This study provides evidence that stereotype overlap indeed can bias relatively low-level perceptual representations of a face in a top-down direction.

Although extrastriate regions along the ventral visual stream are widely acknowledged substrates of static cue processing, with important implications for social categorization, dynamic facial expressions and human movement in general appear to be processed primarily by the superior temporal sulcus (STS), an aspect of the lateral temporal lobe (Haxby et al., 2000). MVPA approaches implicate the STS in representing emotion category information (Said, Moore, Engell, Todorov, & Haxby, 2010b), which largely depends on dynamic facial cues. The STS also has been implicated in fMRI studies of biological motion (Grossman et al., 2000). Ongoing research is attempting to resolve how static and dynamic cues are integrated into stable percepts in the brain. The STS processes and represents both static and dynamic cues but seems to receive this information through divergent pathways. For example, transcranial magnetic stimulation (which temporarily produces effects similar to a focal lesion) to the OFA reduces STS responses to static cues but not to dynamic cues (Pitcher, Duchaine, & Walsh, 2014), suggesting that the STS receives information about static cues from the OFA but receives information about dynamic cues from another, as-yet undescribed pathway.

This is particularly interesting, given the fact that there are no direct white matter pathways from the OFA to the STS (Gschwind, Pourtois, Schwartz, Van De Ville, & Vuilleumier, 2012), further complicating the nature of the relationship. Additional work is necessary to disentangle the processes that integrate static and dynamic cues in person perception.

Researchers interested in trait attribution and impression formation have focused a large amount of their work on the medial PFC (MPFC), which appears to represent the abstract knowledge about personality traits and mental states inferred about another person (Mitchell, Banaji, & Macrae, 2005). This work has isolated an interesting distinction between ventral and dorsal aspects of the MPFC, observing that DMPFC has a more central role in integrating person knowledge with trait attributions gleaned from faces (Ferrari et al., 2016). However, the visual perception of personality traits, as discussed previously, depends on many of the same perceptual mechanisms responsible for representing social category information, and as such, similar regions like the FG also are recruited during trait attributions, such as judgments of trustworthiness (Todorov & Engell, 2008) and baby-facedness (which correlate with judgments of competence/dominance; Zebrowitz, Luevano, Bronstad, & Aharon, 2009). The amygdala, a subcortical structure located deep in the medial temporal lobes, also plays an interesting role tin perceptions of trustworthiness. The amygdala is responsive to trustworthy and untrustworthy faces (Engell, Haxby, & Todorov, 2007), and tracks trustworthiness cues even when faces are presented subliminally (Freeman, Stolier, Ingbretsen, & Hehman, 2014). The amygdala seems preferentially involved in spontaneous trait inference (i.e., rapid personality judgments made without prior knowledge about faces; Todorov & Engell, 2008), consistent with interpretations of amygdala activity as reflecting motivational salience in the environment more generally, signaling relevant details in the visual field and prompting increased processing (W. A. Cunningham & Brosch, 2012). Consistent with recent evidence of the role of facial averageness in trustworthiness judgments (Sofer et al., 2015), the amygdala also has been shown to respond to the atypicality of faces broadly, which may contribute to its observed role in tracking trustworthiness (Said, Dotsch, & Todorov, 2010a).

Although much work in neuroimaging has focused on specific regions that seem selectively tuned to faces, recent work in social and cognitive neuroscience has examined the collaborative role of multiple domain-general neural regions in face perception, with a particular focus on the regions responsible for implementing top-down effects. To characterize a large-scale network involved in top-down face perception, researchers in one study "trained" participants to recognize faces embedded in visual noise. Participants saw faces increasingly obscured by noise and, on critical trials, saw only noise, with no face present. However, participants still reported the ability to detect faces in the pure noise trials (Li et al., 2009). Researchers examined these trials to explore face processing from a completely top-down direction (i.e., when literally no bottom-up facial cues were visually present), and they found greater activity in bilateral FFA and OFA. They additionally examined which regions were most connected functionally with the FFA during illusory face trials and found consistently increased connectivity with the left STS, bilateral OFC, left anterior cingulate cortex (ACC, involved in conflict monitoring; Botvinick, Braver, Barch, Carter, & Cohen, 2001), left dorsolateral prefrontal cortex (DLPFC, involved in response inhibition; MacDonald, Cohen, Stenger, & Carter, 2000), left inferior parietal lobule (IPL), and left premotor cortex (Li et al., 2009). Based on further analysis of both intrinsic and task-based functional connectivity between these regions (Li et al., 2010), the researchers outline a specific model for top-down face processing, wherein the OFC shifts feature processing in the OFA based on predictive assumptions, which can then cause false or biased detection of features in the environment, in turn biasing percepts formed in the FG.

More work is needed to further characterize the top-down face-processing network, but it is interesting to note in the meantime that these neural regions were implicated previously in downstream evaluative aspects of social categorization, such as implicit evaluation (e.g., amygdala, OFC) and regulatory processes exerted on these evaluative biases (e.g., ACC, DLPFC; for a review, see Amodio, 2014). However, additional recent research converges with the idea of a top-down face-processing network, implicating these regions in mechanisms of social categorization when faces simply are viewed passively. A recent neuroimaging study examined neural responses when participants passively viewed face stimuli crossed on race and emotion. Participants also completed a mouse-tracking task to determine individual differences in subjective stereotype overlap between race and emotion categories. As discussed previously, race and emotion categories are linked perceptually because of stereotypic associations (e.g., Black faces are expected to be angry, White faces are expected to be joyful). As faces became more stereotypically incongruent (e.g., joyful Black faces and angry White faces), activity in the ACC increased linearly, and functional connectivity increased between the ACC and FG, reflecting the need to resolve a processing conflict: Due to stereotypes, participants expect to see angry Black faces, and regulatory processing is required to manage the conflict presented by a joyful Black face. Importantly, in individuals high in stereotype overlap, the DLPFC also showed increased activity in response to stereotype-incongruent targets, suggesting a potential inhibition of the initial (incorrect) stereotypical prediction of the face's emotion (Hehman, Ingbretsen, & Freeman, 2014c). The fact that regions involved in conflict monitoring and response inhibition responded when participants were viewing faces passively is an important insight into the large-scale mechanisms of face perception.

Ongoing findings from social neuroscience and social vision have inspired theoretical models that account for the dynamic interactivity between bottom-up and top-down factors observed in person perception. In particular, a computational model has provided a framework for understanding how perceptual cues interact with top-down factors during social categorization and how the precise nature of these interactions can result in many of the perceptual biases reviewed in this chapter (Freeman & Ambady, 2011a). The dynamic interactive (DI) model describes computationally the multiple levels of person perception and how these different levels interact dynamically. At the cue level, the model accounts for visual (face, body, and visual contextual) and auditory (voice) cues, which have ascendant connections to higher-level nodes at the category and stereotype level in addition to feedback connections from those nodes. The category level models social categories, such as sex, race, age, and emotion, which have bidirectional connections to nodes at the stereotype level, allowing stereotypes to provide a dynamic constraint on early-level visual processing of social categories, as we have observed. At each level, nodes are activated in parallel and share lateral connections. For example, individual facial cues can facilitate processing of one another, or they can inhibit one another via feedback connections from stereotype nodes linked to a particular facial feature. The model also accounts for higher-level cognitive states, such as goals, motivations, and attention, which provide an additional constraint on processing. For example, a typical behavioral experiment that tasks a participant with categorizing race activates race-related attentional and goal-oriented processing, which can guide and constrain processing at all levels of the model.

Thus far, the DI model appears to model successfully many of the top-down effects on person perception discussed in this chapter, including complex interactions between perceptual cues and stereotypes, such as intersectionality (Johnson et al., 2012) and overgeneralization effects (Zebrowitz et al., 2003). Recent work also has used the DI model to demonstrate the role of interracial exposure in influencing perceptions of racially ambiguous individuals. Using a large online sample from across the United States, researchers used computer mouse-tracking to show that White participants who live in parts of the country with higher exposure to Black individuals show stable coactivation of the White and Black race categories when categorizing a mixed-race individual. In contrast, White participants who live in areas with low exposure to Black individuals showed highly unstable coactivation of the White and Black categories, with abrupt shifts between the two categories reflected in their computer mouse movements en route to the "White" or "Black" category labels on the screen. Additional modeling work suggested that these unstable dynamics arose in low-exposure individuals because they hold stereotypes that White and Black individuals are highly dissimilar. Thus, when encountering a mixed-race face, bottom-up visual processing attempts to push the system to an in-between point between the White and Black categories, while top-down conceptual knowledge in a low-exposure perceiver tries to rapidly pull the system into a given race category, creating unstable dynamics (Freeman, Pauker, & Sanchez, 2016). Importantly, unstable shifts in race category activation also predicted participants' evaluative bias toward mixed-race individuals, demonstrating that such downstream effects can be represented by individual differences in the structure and dynamics of the broader system (Freeman et al., 2016).

The strength of the DI model may lie in modeling the system as massively interactive, allowing bottom-up perceptual cues, conceptual processing, and high-level cognitive states to interact flexibly in real time during the processing of visual input. Future work is required to refine the model to account for additional cognitive and affective states of the perceiver (e.g., goals, emotional state, attention) and how these inputs affect lower-level perception. To further situate our discussion of person perception in the complex and variable nature of everyday social contexts, we now briefly discuss the impact that certain aspects of the person perception process can have on interpersonal interaction and downstream behavior.

DOWNSTREAM CONSEQUENCES OF PERSON PERCEPTION

Historically, social psychologists studied the downstream consequences of person perception and social categorization while cognitive psychologists and neuroscientists studied the perceptual processes leading to these categorizations and impressions. Much early work in social psychology discussed the consequences of social categorization and stereotyping, recognizing the ability of these processes to guide behavior and evaluative dimensions in an unintended and implicit manner (Bodenhausen & Macrae, 1998). However, this approach assumes a perceptual endpoint: Perceptual processes give rise to an initial impression or categorization, which in turn triggers related stereotypes and their associated biasing power on behavior and interaction. Moreover, researchers also typically assumed that stereotypes were applied to the same degree to all individuals placed in a given social category (Fiske & Taylor, 1991). However, many of the ongoing contributions summarized in this chapter—from interdisciplinary "social vision" approaches to current theoretical models of person perception—suggest that these processes are best thought of as interactive and continuous, and many of the downstream effects of person perception (such as attitudes and biases that can come to pervade society at an institutional level) are rooted in the subtle dynamics of earlier processing stages (Freeman & Ambady, 2011a).

As hinted at in the earlier discussion of facial cues, visual context, and stereotypes, there is a complex relationship between features of the face and their related stereotypes. Consistent with these insights, emerging work shows that attitudes and behaviors emerging from stereotypes and related trait assumptions are not applied categorically to all members of a social category but are often nuanced inferences that depend on the degree to which an individual displays specific category-relevant facial features (Blair, Chapleau, & Judd, 2004a). For example, members of marginalized racial groups experience more stigma and bias when they possess facial cues that are highly typical and diagnostic for their race (Maddox, 2004). In many cases, these biases are enormously consequential for members of marginalized social groups. For example, there exists an unfortunate and well-documented "shooter bias," in which law enforcement officers are more likely to shoot individuals belonging to social categories commonly stereotyped as hostile and aggressive, such as African Americans (Correll, Park, Judd, & Wittenbrink, 2002; but see James, Vila, & Daratha, 2013; James, Klinger, & Vila, 2014). Additionally, researchers found that Black felony offenders with more Afrocentric facial features are given longer criminal sentences, even when controlling for the number and severity of crimes committed and the individuals' past criminal histories (Blair, Judd, & Chapleau, 2004b). These featural biases underscore the importance of studying the impact of bottom-up cues in person perception.

Facial cues also can lead to consequential social outcomes when they introduce inconsistencies and incongruities into social categorization. For example, women with masculine facial features are less likely to be elected to political office in conservative states of the United States (Hehman et al., 2014a). There are also several negative social outcomes for individuals belonging to intersectional social categories (e.g., Black individuals and Asian individuals, because of the association of the "Black" category with masculinity and the "Asian" category with femininity). In a collection of studies, Galinsky, Hall, & Cuddy (2013) found that Black women and Asian men are less successful in heterosexual romantic relationships, and Black individuals in general are more likely to be represented in social institutions that privilege masculinity (such as business leadership and sports). These patterns were found to be consistent between lab-based paradigms and large-scale data from the United States Census and the NCAA Student-Athlete Ethnicity Report (Galinsky et al., 2013).

Rapid trait-based inferences, which depend on invariant facial cues and often are subject to systematic bias, also can result in a number of consequences for the targets of perception. The consequences of perceptions of dominance and competence have been widely studied in the context of leadership selection and compensation. For example, a number of studies show that individuals with faces judged to be competent are more likely to be elected to political office (Ballew & Todorov, 2007). In the business domain, such individuals are much more likely to be hired by successful companies and receive higher salaries (Rule & Ambady, 2008b). Additionally, individuals with untrustworthy-looking faces are more likely to receive guilty verdicts in court, even when there is scant evidence of their guilt (Porter, ten Brinke, & Gustaw, 2010). On the other end of the spectrum, baby-faced individuals (who are consistently perceived as more trustworthy and honest) are more likely to win their legal cases in court (Zebrowitz & McDonald, 1991). Although these insights are discouraging, it is remarkable that early perceptual biases in person perception can come to drive pervasive trends in society at an institutional level. Further work is needed to understand how these insights can be used to implement policies to reduce these systematic trends.

CONCLUSION

The human perceptual system is remarkably attuned to social information in the environment. As such, a particular attentional focus on faces, bodies, and voices emerges early in development and remains a potent force in social perception and interaction throughout the life span. Emerging from the vast cognitive resources deployed to make sense of the social environment is an incredible human propensity to infer quickly and often accurately information from the qualities of another person's face and body, although, as we have discussed, the efficiency of this system sometimes renders it vulnerable to systematic biases. Since interpersonal perception, communication, competition, and cooperation are such integral parts of social life, these biases can be incredibly costly, marring social behavior. Recent theoretical insights, motivated by interdisciplinary collaboration among the social, cognitive, neural, and vision sciences, have provided a deeper understanding and appreciation for how the complexity of the social environment is perceived and understood. The emerging scientific understanding of person perception yields a fascinating picture: a complex and dynamic system encompassing interaction among the perceptual cues available in the environment and the cognitive systems inherent to the perceiver. In this chapter, we have but scratched the surface of this fascinating research domain, and we look forward to future work building an increasingly rich understanding of how we perceive and shape the social world.

REFERENCES

Alaei, R., & Rule, N.O. (2016). Accuracy of perceiving social attributes. In J. A. Hall, M. Schmid Mast, & T. V. West (Eds.), *The Social Psychology of Perceiving Others Accurately* (pp. 125–142). New York, NY: Cambridge University Press.

Allport, G. W. (1954). *The Nature of Prejudice*. Oxford, UK: Addison-Wesley.

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: the importance of facial dynamic in interpreting subtle facial expressions. *Psychological Science*, *16*, 403–410.

Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, *77*, 538–547.

Ambady, N., & Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*(3), 431–441.

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, *15*(10), 670–682.

Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The visual impact of gossip. *Science*, *332*, 1446–1448.

Anzellotti, S., & Caramazza, A. (2014). The neural mechanisms for the recognition of face identity in humans. *Frontiers in Psychology*, *5*, 672.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, *33*, 717–746.

Atkinson, J. W., & Walker, E. L. (1956). The affiliation motive and perceptual sensitivity to faces. *Journal of Abnormal Psychology*, *53*, 38–41.

Balcetis, E., & Dunning, D. (2006). See what you want to see: motivational influences on visual perception. *Journal of Personality and Social Psychology*, *91*, 612–625.

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*, 17948–17953.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*, 617–629.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, *20*, 286–290.

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, *23*, 361–372.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*, 217–239.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: the impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, *100*, 3–18.

Bestelmeyer, P. E. G., Rouger, J., DeBruine, L. M., & Belin, P. (2010). Auditory adaptation in vocal affect perception. *Cognition*, *117*, 217–223.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*, 115–147.

Blair, I. V., Chapleau, K. M., & Judd, C. M. (2004a). The use of Afrocentric features as cues for judgment in the presence of diagnostic information. *European Journal of Social Psychology*, *34*, 1–10.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004b). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*, 674 – 679.

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind*, *2*, 5–38.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In J. R. Wyer (Ed.), *Advances in Social Cognition* (Vol. 11, pp. 1–52). Mahwah, NJ: Erlbaum.

Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Brewer, M. B. (1988). A dual process model of impression formation. *Advances in Social Cognition*, *1*, 1–36.

Brinsmead-Stockham, K., Johnston, L., Miles, L., & Macrae, C. N. (2008). Female sexual orientation and menstrual influences on person perception. *Journal of Experimental Social Psychology*, *44*, 729–734.

Brosch, T., Bar-David, E., & Phelps, E. A. (2013). Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychological Science*, *24*(2), 160–166.

Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception*, *22*, 829–840.

Bruce, V., & Young, A. W. (1986). A theoretical perspective for understanding face recognition. *British Journal of Psychology*, *77*, 305–327.

Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, *42*, 33–44.

Bullier, J. (2002). Neural basis of vision. In H. Pashler & S. Yantis (Eds.), *Stevens' Handbook of Experimental Psychology, Vol. 1: Sensation and Perception* (pp. 1–40). Hoboken, NJ: Wiley.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*, 535−543.

Carré, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2651–2656.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science*, *20*, 1194–1198.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*, *70*, 205–218.

Chiao, J., Heck, H. E., Nakayama, K., & Ambady, N. (2006). Priming race in biracial observers affects visual search for Black and White faces. *Psychological Science*, *17*, 387–392.

Contreras, J. M., Banaji, M. R., & Mitchell, J. P. (2013). Multivoxel patterns in fusiform face area differentiate faces by sex and race. *PLoS ONE*, *8*(7), e69684.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*(6), 1314–1329.

Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*(5667), 115–117.

Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology*, *50*, 925–935.

Cunningham, W. A., & Brosch, T. (2012). Motivational salience: amygdala tuning from traits, needs, values, and goals. *Current Directions in Psychological Science*, *21*(1), 54–59.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, *9*, 353–356.

Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 357–372.

de Gelder, B. (2005). Toward a neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*, 242–249.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*, 289−311.

Downing, P. E., Bray, D., Rogers, J., & Childs, C. (2004). Bodies capture attention when nothing is expected. *Cognition*, *93*, B27–B38.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: race, crime, and visual processing. *Journal of Personality and Social Psychology*, *87*, 876–893.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic encoding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, *19*(9), 1508–1519.

Erdelyi, M. H. (1974). A new look at the new look: perceptual defense and vigilance. *Psychological Review*, *81*(1), 1–25.

Evans, S., Neave, N., & Wakelin, D. (2006). Relationships between vocal characteristics and body size and shape in human males: an evolutionary explanation for a deep male voice. *Biological Psychology*, *72*, 160–163.

Ferrari, C., Lega, C., Vernice, M., Tamietto, M., Mende-Siedlecki, P., Vecchi, T., Todorov, A., & Cattaneo, Z. (2014). The dorsomedial prefrontal cortex plays a causal role in integrating social impressions from faces and verbal descriptions. *Cerebral Cortex*, *26*(1), 156–165.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*: e229.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*, 77–83.

Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition*, 2nd ed. New York, NY: McGraw-Hill.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, *106*, 1511–1522.

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Fraley, R. C., Niedenthal, P. M., Marks, M. J., Brumbaugh, C. C., & Vicary, A. (2006). Adult attachment and the perception of emotional expressions: probing the hyper-activating strategies underlying anxious attachment. *Journal of Personality*, *74*, 1163–1190.

Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, *20*, 1183–1188.

Freeman, J. B., & Ambady, N. (2011a). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279.

Freeman, J. B., & Ambady, N. (2011b). Hand movements reveal the time-course of shape and pigmentation processing in face categorization. *Psychonomic Bulletin & Review*, *18*, 705–712.

Freeman, J. B., & Ambady, N. (2011c). When two become one: temporally dynamic integration of the face and voice. *Journal of Experimental Social Psychology*, *47*, 259–263.

Freeman, J. B., Johnson, K. L., Ambady, N., & Rule, N. O. (2010). Sexual orientation perception involves gendered facial cues. *Personality and Social Psychology Bulletin*, *36*(10), 1318–1331.

Freeman, J. B., Ma, Y., Han, S., & Ambady, N. (2013). Influences of culture and visual context on real-time social categorization. *Journal of Experimental Social Psychology*, *49*, 206–210.

Freeman, J. B., Pauker, K., & Sanchez, D. T. (2016). A perceptual pathway to bias: interracial exposure reduces abrupt shifts in real-time race perception that predict mixed-race bias. *Psychological Science*, *27*(4), 502–517.

Freeman, J. B., Penner, A. M., Saperstein, A., Scheutz, M., & Ambady, N. (2011). Looking the part: social status cues shape race perception. *PLoS ONE*, *6*, e25107.

Freeman, J. B., Stolier, R. M., Ingbretsen, Z. A., & Hehman, E. A. (2014). Amygdala responsivity to high-level social information from unseen faces. *Journal of Neuroscience*, *34*(32), 10573–10581.

Freire, A., Lewis, T. L., Maurer, D., & Blake, R. (2006). The development of sensitivity to biological motion in noise. *Perception*, *35*(5), 647–657.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694–724.

Galinsky, A. D., Hall, E. V., & Cuddy, A. J. (2013). Gendered races: implications for interracial marriage, leadership selection, and athletic participation. *Psychological Science*, *24*(4), 498–506.

Gangestad, S. W., Simpson, J. A., DiGeronimo, K., & Biek, M. (1992). Differential accuracy in person perception across traits: examination of a functional hypothesis. *Journal of Personality and Social Psychology*, *62*, 688–698.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Gilbert, D.T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4th ed.). New York, NY: McGraw Hill.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*, 32–41.

Goldberg, L. R. (1990). An alternate "description of personality": the big-five factor structure. *Personality Processes and Individual Differences*, *59*(6), 1216–1229.

Goodale, M. A., & Milner, A. D. (1992). Separate pathways for perception and action. *Trends in Neurosciences*, *15*, 20–25.

Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: the role of symmetry and averageness. *Journal of Comparative Psychology*, *108*, 233–242.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*(8), 536–548.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*(5), 711–720.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2012). White-matter connectivity between face-responsive regions in the human brain. *Cerebral Cortex*, *22*(7), 1564–1576.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*, 330–337.

Harrigan, J. A., Harrigan, K. M., Sale, B. A., & Rosenthal, R. (1996). Detecting anxiety and defensiveness from visual and auditory cues. *Journal of Personality*, *64*, 675–709.

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 571–576.

Hassin, R., Bargh, J. A., & Uleman, J. S. (2002). Spontaneous causal inferences. *Journal of Experimental Social Psychology*, *38*, 515–522.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. A. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. A. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*, 59–67.

Hehman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. (2014a). Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychology and Personality Science*, *5*, 815–824.

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially impact the consistency of social evaluations. *Personality and Social Psychology Bulletin*, *41*, 1123–1143.

Hehman, E., Ingbretsen, Z. A., & Freeman, J. B. (2014c). The neural basis of stereotypic impact on multiple social categorization. *Psychological Science*, *15*, 342–345.

Hehman, E., Leitner, J. B., & Gaertner, S. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, *49*, 747–754.

Hehman, E., Leitner, J. B., & Freeman, J. B. (2014b). The face-time continuum: lifespan changes in facial width-to-height ratio impact aging-associated perceptions. *Personality and Social Psychology Bulletin*, *40*, 1624–1636.

Helmholtz, H. (1867). Concerning the perception in general. In *Readings in the History of Psychology*. East Norfolk, CT: Appleton-Century-Crofts.

Hess, U., Adams, R. B., & Kleck, R. E. (2004). Facial appearance, gender, and emotion expression. *Emotion*, *4*, 378–388.

Hess, U., Adams, R. B., & Kleck, R. E. (2005). Who may frown and who should smile? Dominance, affiliation, and the display of happiness and anger. *Cognition & Emotion*, *19*, 515–536.

Hess, U., Senécal, S., Kirouac, G., Herrera, P., Philippot, P., & Kleck, R. E. (2000). Emotional expressivity in men and women: Stereotypes and self-perceptions. *Cognition & Emotion*, *14*(5), 609–642.

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: the role of shape and colour. *Proceedings of the Royal Society B*, *261*, 367–373.

Hills, P. J., & Lewis, M. B. (2006). Reducing the own-race bias in face recognition by shifting attention. *Quarterly Journal of Experimental Psychology*, *59*, 996–1002.

Hills, P. J., & Lewis, M. B. (2011). Reducing the own-race bias in face recognition by attentional shift using fixation crosses preceding the lower half of a face. *Visual Cognition*, *19*, 313–339.

Hills, P. J., & Pake, J. M. (2013). Eye-tracking the own-race bias in face recognition: revealing the perceptual and socio-cognitive mechanisms. *Cognition*, *129*, 586–597.

Hugenberg, K., & Bodenhausen, G.V. (2004). Ambiguity in social categorization: the role of prejudice and facial affect in race categorization. *Psychological Science*, *15*, 342–345.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187.

Hughes, S. M., & Rhodes, G. (2010) Making age assessments based on voice: the impact of the reproductive viability of the speaker. *Journal of Social, Evolutionary, & Cultural Psychology*, *4*, 290–304.

Hutchings, P. B., & Haddock, G. (2008). Look black in anger: the role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology*, *44*, 1418–1420.

Inzlicht, M., Kaiser, C. R., & Major, B. (2008). The face of chauvinism: how prejudice expectations shape perceptions of facial affect. *Journal of Experimental Social Psychology*, *44*, 758–766.

Ito, K., Masuda, T., & Hioki, K. (2012). Affective information in context and judgment of facial expression: cultural similarities and variations in context effects between North Americans and East Asians. *Journal of Cross-Cultural Psychology*, *43*, 429–445.

James, L., Klinger, D. & Vila, B. (2014). Racial and ethnic bias in decisions to shoot seen through a stronger lens: experimental results from high-fidelity laboratory simulations. *Journal of Experimental Criminology*, *10*, 323–340.

James, L., Vila, B., & Daratha, K. (2013). Influence of suspect race and ethnicity on decisions to shoot in high fidelity deadly force judgment and decision-making simulations. *Journal of Experimental Criminology*, *9*, 189–212.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201–211.

Johansson, G. (1975). Visual motion perception. *Scientific American*, *232*, 76–89.

Johnson, K. L., & Freeman, J. B. (2010). A “new look” at person construal: seeing beyond dominance and discreteness. In E. Balcetis & D. Lassiter (Eds.), *The Social Psychology of Visual Perception* (pp. 253–272). New York, NY: Psychology Press.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: how covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, *102*, 116–131.

Johnson, K. L., Gill, S., Reichman, V., & Tassinary, L. G. (2007). Swagger, sway, and sexuality: judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology*, *93*, 321–334.

Johnston, L., Arden, K., Macrae, C. N., & Grace, R. C. (2003). The need for speed: the menstrual cycle and person construal. *Social Cognition*, *21*, 89–99.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Kaul, C., Rees, G., & Ishai, A. (2011). The gender of face stimuli is represented in multiple regions in the human brain. *Frontiers in Human Neuroscience*, *4*, 238.

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, *98*, 155–163.

Kluver, M., Hecht, H., & Troje, N. F. (2016). Internal consistency predicts attractiveness in biological motion walkers. *Evolution and Human Behavior*, *37*(1), 40–46.

Knutson, K. M., Mah, L., Manly, C. F., & Grafman, J. (2007). Neural correlates of automatic beliefs about gender and race. *Human Brain Mapping*, *28*(10), 915–930.

Ko, S. J., Judd, C. M., & Blair, I. V. (2006). What the voice reveals: within- and between-category stereotyping on the basis of voice. *Personality and Social Psychology Bulletin*, *32*, 806–819.

Kozlowski, L.T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, *21*, 575–580.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: a parallel-constraint-satisfaction theory. *Psychological Review*, *103*, 284 –308.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(26), 15387–15392.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallamm, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*, 390–423.

Langlois, J. H., & Roggman, L.A. (1990). Attractive faces are only average. *Psychological Science*, *1*, 115–121.

Latinus, M., & Belin, P. (2012). Perceptual auditory aftereffects on voice identification using brief vowel stimuli. *PLoS ONE*, *7*, e41384.

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: facial width is associated with testosterone levels in men. *Evolution & Human Behavior*, *34*, 273–279.

Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces: the role of race categories. *Journal of Experimental Psychology: General*, *135*(4), 501–512.

Li, J., Liu, J., Liang, J., Zhang, H., Zhao, J., Huber, D. E., . . . Shi, G. (2009). A distributed neural system for top-down face processing. *Neuroscience Letters*, *451*(1), 6–10.

Li, J., Liu, J., Liang, J., Zhang, H., Zhao, J., Rieth, C. A., . . . Lee, K. (2010). Effective connectivities of cortical regions for top-down face processing: a dynamic causal modeling study. *Brain Research*, *1340*, 40–51.

Little, A. C., Burt, D. M., Penton-Voak, I. S., & Parrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry of male faces. *Proceedings of the Royal Society of London B*, *268*, 39–44.

Macchi, C. V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science*, *15*(6), 379–383.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *228*(5472), 1835–1838.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: the ambiguous-race face effect. *Psychology, Public Policy and Law*, *7*, 98–118.

Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review*, *8*(4), 383–401.

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology*, *94*, 365–381.

Masuda, S., Tsujii, T., & Watanabe, S. (2005). An interference effect of voice presentation on face gender discrimination task: evidence from event-related potentials. *International Congress Series*, *1278*, 156−159.

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say "hello"? Personality impressions from brief novel voices. *PLoS ONE*, *9*(3): e90779.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*, 215–238.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy, and Law*, *7*, 3–35.

Michel, C., Corneille, O., & Roisson, B. (2007). Race categorization modulated holistic face processing. *Cognitive Science*, *31*, 911–924.

Michel, C., Rossion, B., Han, J., Chung, C., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, *17*, 608–615.

Milne, E., & Grafman, J. (2001). Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *Journal of Neuroscience*, *21*(12): RC150.

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, *28*, 757–762.

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*, 237–254.

Montepare, J. M., & Zebrowitz, L. A. (1993). A cross-cultural comparison of impressions created by age-related variations in gait. *Journal of Nonverbal Behavior*, *17*, 55–68.

Montepare, J. M., & Zebrowitz, L. A. (1998). "Person perception comes of age": the salience and significance of age in social judgments. In M. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 30, pp. 93–163). San Diego, CA: Academic Press.

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*, 3–18.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*, 424–430.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 11087–11092.

Ostrom, T. M., & Sedikides, C. (1992). Out-group homogeneity effects in natural and minimal groups. *Psychological Bulletin*, *112*(3), 536–552.

Pauker, K., & Ambady, N. (2009). Multiracial faces: how categorization affects memory at the boundaries of race. *Journal of Social Issues*, *65*, 69–86.

Pauker, K., Rule, N. O., & Ambady, N. (2010). Ambiguity and social perception. In E. Balcetis & D. Lassiter (Eds.), *The Social Psychology of Visual Perception* (pp. 7–26). New York, NY: Psychology Press.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*, 636–648.

Pinel, E. C. (1999). Stigma consciousness: the psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, *76*(1), 114–128.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology*, *24*(17), 2066–2070.

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: the impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime, and Law*, *16*, 477–491.

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology*, *41A*, 797–828.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365.

Ratner, K. G., Kaul, C., & Van Bavel, J. J. (2013). Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. *Social Cognitive and Affective Neuroscience*, *8*, 750–755.

Re, D. E., & Rule, N. O. (2016). Predicting firm success from the facial appearance of chief executive officers of non-profit organizations. *Perception*, *45*, 1137–1150.

Reed, C. L., Stone, V., Bozova, S., & Tanaka, J. (2003). The body inversion effect. *Psychological Science*, *14*, 302–308.

Richardson, M. J., & Johnston, L. (2005). Person recognition from dynamic events: the kinematic specification of individual identity in walking style. *Journal of Nonverbal Behavior*, *29*, 25–44.

Richeson, J. A., & Shelton, J. N. (2005). Thin slices of racial bias. *Journal of Nonverbal Behavior*, *29*, 211–236.

Righart, R., & de Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Social Cognitive and Affective Neuroscience*, *3*, 270–278.

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience*, *8*, 107–113.

Rule, N. O., & Ambady, N. (2008a). Brief exposures: male sexual orientation is accurately perceived at 50ms. *Journal of Experimental Social Psychology*, *44*, 1100–1105.

Rule, N. O., & Ambady, N. (2008b). The face of success: inferences from chief executive officers' appearance predict company profits. *Psychological Science*, *19*, 109–111.

Rule, N. O., Ambady, N., Adams, R. B., & Macrae, C. N. (2008). Accuracy and awareness in the perception and categorization of male sexual orientation. *Journal of Personality and Social Psychology*, *95*, 1019–1028.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*, 585–615.

Said, C. P., Dotsch, R., & Todorov, A. (2010a). The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia.* *48*(12), 3596–3605.

Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010b). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, *10*(5), 11.

Savic, I., Berglund, H., Gulyas, B., & Roland, P. (2001). Smelling of odorous sex hormone–like compounds causes sex-differentiated hypothalamic activations in humans. *Neuron*, *31*, 661–668.

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry, and cues to food genes. *Proceedings of the Royal Society of London B*, *266*, 1913–1917.

Scherer, K R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Communication*, *40*, 227–256.

Simion, F., & Di Giorgio, E. (2015). Face perception and processing in early infancy: inborn predispositions and developmental changes. *Frontiers in Psychology*, *6*(969).

Smith, E. L., Grabowecky, M., & Suzuki, S. (2007). Auditory–visual crossmodal integration in perception of face gender. *Current Biology*, *17*, 1680−1685.

Sofer, C., Dotsch, R., Wigboldus, D. H. J., & Todorov, A. (2015). What is typical is good: the influence of face typicality on perceived trustworthiness. *Psychological Science*, *26*(1), 39–47.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychological Science*, *21*, 349–354.

Stolier, R. M., & Freeman, J. B. (2016). Neural pattern similarity reveals the inherent intersection of social categories. *Nature Neuroscience*, *19*, 795–797.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409.

Sutherland, C. A. M., Rowley, L. E., Amoaku, U. T., Daguzan, E., Kidd-Rossiter, K. A., Maceviciute, U., & Young, A. W. (2015). Personality judgments from everyday images of faces. *Frontiers in Psychology*, *6*(1616), 1–11.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96–102.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: averageness, symmetry, and parasite resistance. *Human Nature*, *4*, 237–269.

Todorov, A. (2011). Evaluating faces on social dimensions. In A. Todorov, S. T. Fiske, D. A. Prentice (Eds.), *Social Neuroscience: Toward Understanding the Underpinnings of the Social Mind* (pp. 54–76). New York, NY: Oxford University Press.

Todorov, A., & Engell, A. D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social Cognitive and Affective Neuroscience*, *3*(4), 303–312.

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*, 163–173.

Troje, N. F. (2003). Gender and attractiveness from biological motion. *Journal of Vision*, *3*(9), 86a.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Uleman, J. S., Blader, S., & Todorov, A. (2005). Implicit impressions. In R. Hassin, J. S. Uleman, & J. A. Barge (Eds.), *The New Unconscious* (pp. 362–392). New York, NY: Oxford University Press.

Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin*, *35*, 321–335.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychological Science*, *19*(11), 1131–1139.

Van Bavel, J. J., Swencionis, J. O'Connor, R., & Cunningham, W. A. (2012). Motivated social memory: belonging needs moderate the own-group bias in face recognition. *Journal of Experimental Social Psychology*, *48*, 707–713.

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*, 487–494.

Van Dommelen, W. A. (1993). Speaker height and weight identification: a re-evaluation of some old data. *Journal of Phonetics*, *21*, 337–341.

Voss, A., Rothermund, K., & Brandtstadter, J. (2008). Interpreting ambiguous stimuli: separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*, 1048–1056.

Walton, J. H., & Orlikoff, R. F. (1994). Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech, Language, and Hearing Research*, *37*(4), 738–745.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology*, *57*, 120–128.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural face categories. *Nature*, *428*, 557–561.

Weisbuch, M., & Ambady, N. (2011). Thin-slice vision. In R. B. Adams, N. Ambady, K. Nakayama, & S. Shimojo, *The Science of Social Vision* (pp. 228–247). New York, NY: Oxford University Press.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience*, *14*(5), 656–661.

Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: evidence from connectionist modeling. *Personality and Social Psychology Review*, *7*, 194–215.

Zebrowitz, L. A., Luevano, V. X., Bronstad, P. M., & Aharon, I. (2009). Neural activation to babyfaced men matches activation to babies. *Social Neuroscience*, *4*(1), 1–10.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, *15*(6), 603–623.

Zebrowitz, L. A., & Montepare, J. M. (2006). The ecological approach to person perception: evolutionary roots and contemporary offshoots. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and Social Psychology* (pp. 81–113). New York, NY: Psychology Press.

CHAPTER 14

Group Processes and Intergroup Relations

MICHAEL A. HOGG AND AMBER M. GAFFNEY

GROUP PROCESSES AND INTERGROUP RELATIONS

Try to imagine a world without groups. It is almost impossible. Nearly everything we do and experience involves groups: we work in groups, play in groups, and are brought up in groups. Without groups, there would effectively be no agriculture, no economy, no culture, no religion, no pyramids, no cities, no computers, and so forth; and of course no group membership–based prejudice, discrimination, disadvantage, and genocide. Groups are neither good nor bad; they just are, and they pervade almost all aspects of the human experience. Humans are not merely social animals; we are group animals.

Social psychologists have long studied groups, traditionally from two different perspectives. The study of group dynamics examines the interaction among people in small, face-to-face interactive task-oriented groups (e.g., Forsyth, 2006), and the study of intergroup relations examines how people and groups perceive and interact with or in the context of other groups (e.g., Hogg, 2013). However, since the late 1980s, these two strands have grown increasingly close and conceptually interwoven so that the distinction is now largely blurred in contemporary research on groups (e.g., Levine, 2013; Levine & Hogg, 2010; Stangor, 2016; also see the journal *Group Processes and Intergroup Relations*).

Given this historical legacy and delineation of research, our account of group processes and intergroup relations largely follows the conventional sequence of moving from a discussion primarily of interaction and processes *within groups* to a discussion primarily of interaction and processes *between groups*.

WHAT ARE GROUPS, AND WHY DO PEOPLE JOIN THEM?

What Is a Social Group?

There are almost as many definitions of a group as there are theories of groups in social psychology. However, on one thing social psychologists agree: A lone individual, disconnected from others, is not a group. Social psychologists also largely agree that two people do not really constitute a group—their dynamic is dominated by interpersonal relationship-based processes, not group-based processes. So, the first component of a definition is that a group comprises at least three people. Resting on this common ground, definitions fall into two broad categories: (a) those that emphasize social interaction and relationships among people, and (b) those that emphasize self-conception and cognitive representations.

The social interaction perspective traditionally has framed the group dynamics tradition alluded to in our opening section (cf. Forsyth, 2006). A group is a relatively small collection of individuals who are able to interact, typically face-to-face. We are all familiar with these groups—teams, decision-making groups, work groups, organizations, and others. Members tend to have shared goals and are interdependent in their pursuit, conform to others' behavior, and develop bonds of attraction to one another due to mutual goal satisfaction (e.g., Festinger, Schachter, & Back, 1950; Sherif, 1966; cf. Hogg, 1993).

In contrast, the more cognitive self-conception perspective plays a significant role in the more contemporary analysis of group processes, already alluded to, that is independent of group size and integrates what happens within and what happens between groups (e.g., Levine, 2013; Levine & Hogg, 2010; Stangor, 2016). A group is a collection of individuals who categorize themselves as belonging to the same social category and thus embody the category's attributes and feel an associated sense of attachment to and identification with the category (Turner, Hogg, Oakes, Reicher & Wetherell, 1987; also see Abrams & Hogg, 2010; Hogg, 2006).

From this largely social identity theory perspective, a human group is a social category that people cognitively represent as a prototype—a fuzzy set of attributes (attitudes, behaviors) that capture similarities within the group and differences between the in-group and relevant out-groups. Prototypes tend to accentuate intragroup similarities and intergroup differences, and they tend to be shared among members of the same group (e.g., "we agree that we are like this"); they can be viewed as the cognitive representations of group norms (see Hogg, 2010). We are all familiar with this aspect of groups—belonging to a political party or religion, identifying with a nation or ethnic group, feeling part of an organization or team.

How Groups Differ

Groups and their associated social identities are not all the same. Aside from specific differences in terms of what a group does and what it believes (Democrats versus Republicans, terrorist organization versus Jesuits), there are some more general dimensions that differentiate between groups. One key dimension is entitativity (which is defined as that structural property of a group, derived from clear intergroup boundaries and clear internal structure, and similarity, interdependence, and common fate among members) that makes a group appear as a coherent and distinct entity (Campbell, 1958; Hamilton & Sherman, 1996).

Lickel and colleagues (2000) identify four types of groups that vary from most to least entitative. *Intimacy groups* (e.g., familial groups, friendship groups) are tightly knit groups, characterized by low permeability, high investment and commitment, and enduring membership. *Task groups* are a collection of people not bound by interpersonal relationships but rather by a common goal of completing a task or project (e.g., work groups, student project groups). These groups might disintegrate upon task completion and are easy for members to join or leave. *Social categories* (e.g., gender, ethnic, political groups) are characterized by low interaction but nonetheless provide a source of identification for members, often through a sense of shared history. *Loose associations* (e.g., people in the same neighborhood, Bob Dylan fans) have a high degree of permeability and low interaction among members.

Highly entitative groups can provide a sense of belonging and a clear sense of identity in society (Hogg, 2007, 2012b) but also can be somewhat suffocating and extreme—requiring absolute conformity and rejection of deviants, having hierarchical leadership and polarized norms and intergroup perceptions, and engaging in extreme intergroup behaviors (Hogg, 2014; Kruglanski, Pierro, Mannetti, & De Grada, 2006; Yzerbyt, Judd, & Corneille, 2004).

Another general group difference is between common bond and common identity groups (Prentice, Miller, & Lightdale, 1994). *Common bond* groups are based on bonds of attraction between individual group members—such groups often arise out of shared interests, mutual friendships, or familial ties and disintegrate due to interpersonal conflict among group members. When a member leaves a common bond group, the group's internal structure may collapse and the group may disintegrate. *Common identity* groups are based on shared category membership (e.g., gender, nationality)—such groups may arise out of a shared interest or a common goal (e.g., political party, student organization) or are ascribed to capture shared category attributes of its members (e.g., Asian, woman). When a member leaves a common identity group, the existence of the group and the identity that it provides tend to remain intact.

Related to this distinction between common bond and common identity groups is a distinction among types of group membership–based identity: social identity (Brewer & Gardner, 1996; Chen, Boucher, & Tapias, 2006). For the purposes of this chapter, the key distinction is between social identities that group members define in terms of shared category attributes (collective self, collective identity) and social identities that members define in terms of the network of connections and role relationships among members of a category (relational self, relational social identity).

This distinction may map onto the familiar distinction between individualistic and collectivistic cultures (Hofstede, 2001) and independent and interdependent self-construal (Markus & Kitayama, 2001; also see Yap, Ji, & Hong, 2017). Groups and social identities tend to be organized around the collective self among people from individualistic cultures and/or those who have an independent self-construal and organized around the relational self among people from collectivist cultures and/or those who have an interdependent self-construal (Brewer & Chen, 2007).

Motivations for Affiliation and Group Membership

People affiliate with other people and join and identify with groups for many different reasons. Some very concrete reasons are self-protection and survival and to accomplish something that cannot be accomplished alone (bringing down a mammoth, building a pyramid, performing a symphony); simply being with others can even reduce anxiety (Schachter, 1959).

At a more basic psychological level, humans are a fundamentally social species—we have such an overwhelming need to belong (Baumeister & Leary, 1995) that our whole sense of well-being and self-esteem may rest on it (Leary, Tambor, Terdal, & Downs, 1995), and social isolation, exclusion, and ostracism can have devastating psychological consequences (Abrams, Marques, & Hogg, 2005; Kurzban & Leary, 2001; Williams, 2001). According to terror management theory, thoughts of our own mortality can drive us to endorse ideologies and worldviews that are widely shared and to some extent provide us with a sense of symbolic immortality and collective continuity after death (e.g., Greenberg, Solomon, & Pyszczynski, 1997; Solomon, Greenberg, & Pyszczynski, 1991).

Ultimately our attitudes and worldview, even speech and language, depend on human affiliation and social interaction largely with people who are similar to ourselves—we gravitate toward similar others, toward our in-groups, to make social comparisons and ultimately know what to think, how to behave, and even who we are (Festinger, 1954; Hogg & Gaffney, 2014). This last point underscores the fact that groups define and validate who we are in society and thus configure our attitudes and behaviors and our perceptions and expectations of others. The self-categorization process underlying group identification assigns group attributes and their societal evaluation to the self (Tajfel & Turner, 1986; Turner et al., 1987; see Abrams & Hogg, 2010). Associated with this process are at least three social identity and group membership motivations, described by: (1) the self-esteem hypothesis, (2) the uncertainty-identity theory, and (3) the optimal distinctiveness theory.

The self-esteem hypothesis (Abrams & Hogg, 1988; Rubin & Hewstone, 1998) invokes the pursuit and protection of self-esteem as a fundamental human motivation (Sedikides & Strube, 1997). It argues that in social identity (group) contexts, people are motivated to protect and promote the evaluatively positive distinctiveness of their in-group (from relevant and/or competing out-groups) because by so doing it reflects positively on social identity, self-evaluation, and self-esteem. People are motivated to identify with favorable groups and feel good about themselves if they succeed. However, people are resilient, and membership in unfavorable stigmatized groups does not necessarily translate into low self-esteem (cf. Crocker & Major, 1989).

Uncertainty-identity theory argues that people are fundamentally motivated to reduce feelings of uncertainty about themselves and the world they live in and that the process of social categorization of self and others that underlies group identification and membership is a powerful resolution of this uncertainty. Simply put, people identify with groups to reduce self and identity uncertainty (Hogg, 2007, 2012b). Some types of groups are better at resolving uncertainty than other groups, thus group entitativity moderates the relationship between uncertainty and group identification. Under uncertainty, people prefer to identify with highly entitative groups because such groups provide a clearer and less ambiguous identity than low-entitativity groups. Taken to extremes, uncertainty-sponsored identification with highly entitative groups can manifest as zealotry and societal extremism (Hogg, 2014), which can be associated with group centrism (Kruglanski et al., 2006), political and religious fundamentalism (Hogg, 2015b; Hogg, Adelman, & Blagg, 2010), and autocratic leadership (Hogg, 2005).

Finally, optimal distinctiveness theory (Brewer, 1991; Leonardelli, Pickett, & Brewer, 2010) argues that two competing primary motivations affect group and identity processes: the desire to be included and to be similar to others (assimilation) and the desire to be unique and distinct from other people (differentiation). Because these are contrasting human motives, the equilibrium state is one of optimal distinctiveness. Group membership satisfies optimal distinctiveness because groups provide members with similar (in-group) others and a sense of assimilation and belonging while at the same time providing a sense of differentiation and distinctiveness from out-groups. Typically very small groups oversatisfy the need for differentiation and undersatisfy the need for assimilation, while very large groups do the opposite—there is often a context-dependent optimal group size.

HOW BEING IN THE PHYSICAL PRESENCE OF OTHERS AFFECTS OUR BEHAVIOR

We have some very basic psychological needs that groups fulfill; however, groups also serve a distinct role in developing and changing our attitudes and behavior. Sometimes the mere presence of other people is enough to affect our behavior in both positive and negative ways. Perhaps, then, the most basic question we can ask about groups is how our behavior is affected by being in the presence of other people. There are three main lines of research that address this question: (1) social facilitation, (2) deindividuation, and (3) social loafing.

Social Facilitation

Think about occasions when you have prepared and rehearsed a talk and then stood up in front of an audience to deliver the talk—sometimes (you feel) it went very well, and other times (you feel) it was a disaster. One of the very first experiments in social psychology addressed this issue (Triplett, 1898). Norman Triplett was an avid cyclist who designed an experiment to illustrate what he had already observed: that cyclists often ride harder and faster in the presence of other cyclists than when cycling alone. Although Triplett did not provide an adequate answer to the question why?, subsequent research has done so.

According to drive theory of social facilitation (Zajonc, 1965), the mere physical presence of other people is physiologically and psychologically arousing (one needs to be in a state of readiness when others are around) and serves as a drive that facilitates dominant behavioral responses. On well-practiced tasks, the dominant response is to get it correct, so mere presence drives/facilitates improved performance; on poorly practiced tasks, the dominant response is to get it wrong, so mere presence drives/facilitates impaired performance—people "choke."

Interestingly, research has shown that, contrary to what you might think, it is being in the presence of other people, not apprehension about how they might evaluate you, that best explains social facilitation. To demonstrate this, Schmitt, Gilovich, Goore, and Joseph (1986) had people type their name into a computer (easy) and then type their name backward interspersed with ascending digits (difficult). They did this alone after the experimenter had left the room, in the mere presence of a blindfolded confederate wearing a headset and participating in a different experiment, or under the close observation of the experimenter who remained in the room carefully scrutinizing the participant's performance. Mere presence improved performance of the easy task and impaired performance of the difficult task—evaluation apprehension had little additional impact.

Other explanations of social facilitation focus on the attentional effects of mere presence (e.g., Baron, 1986; Manstead & Semin, 1980). People have limited attentional capacity, and the mere presence of other people is distracting—it redirects attention from the task at hand to the audience. If the task is poorly learned and thus difficult, people need to focus all of their attention on it. The presence of an audience limits the amount of attention they can focus on the task, so performance deteriorates. If the task is well learned and thus easy, people typically do not pay much attention to it. The presence of an audience limits the amount of attention people can focus on the task; so they concentrate more on the task and thus performance improves.

Deindividuation

Being in the presence of other people, particularly in a large crowd, has other effects. According to early crowd theorists (e.g., LeBon, 1908), people in crowds descend several rungs on the ladder of civilization to become unsocialized instinct-driven barbarians. Crowds produce this behavior because people feel anonymous and devoid of personal responsibility for their actions, which encourages the release of unconscious antisocial motives ("ancestral savagery") through suggestion (a process akin to hypnosis) and allows ideas and sentiments to spread rapidly and unpredictably through a process of contagion.

Zimbardo (1969) gave this idea a modern spin that focuses on how anonymity and lack of personal accountability create a psychological state of deindividuation that "liberates" people from responsibility and the constraints of socialized norms, and allows them to engage in selfish, antisocial, and often aggressive behaviors. Although deindividuation can arise through a sense of being lost in the crowd, even isolated individuals can behave in a deindividuated manner if they feel anonymous and unaccountable.

Traditionally, deindividuation is about loss of identity. In contrast, Reicher, Spears, and Postmes (1995) have proposed that what actually happens in crowds and other collective situations is that people experience a transformation of identity—rather than losing identity, they recategorize themselves in terms of a situated social identity (e.g., football supporters, civil rights protesters, Megadeath fans) that prescribes and regulates exactly how to behave in that situation. There is a logic to crowd behavior, and in most cases there is a clear intergroup dimension—for example, protesters versus police, Real Madrid fans versus Atlético de Madrid fans.

Social Loafing

The deindividuation premise that people in groups are less than they could be is part of a wider assumption. Even if people in groups do not embark on a narcissistic orgy of antisocial and aggressive behavior, they certainly do tend to exert less effort on a task and sit back and let others carry the load—they socially loaf. Social loafing is a tendency for people to work less hard (i.e., loaf) on a task when they believe others are also working on the task; there is "a reduction in individual effort when working on a collective task (in which one's outputs are pooled with those of other group members) compared to when working either alone or coactively" (Williams, Karau, & Bourgeois, 1993, p. 131).

A meta-analytic review by Karau and Williams (1993) of the 78 social loafing studies conducted up to the early 1990s found loafing to be widespread and common. It has been found in the laboratory as well as in the field and in both Western and Asian cultures. The effect has been recorded using physical tasks (e.g., clapping, rope pulling, and swimming), cognitive tasks (e.g., generating ideas), evaluative tasks (e.g., quality ratings of poems) and perceptual tasks (e.g., maze performance). People even loaf when tipping in restaurants!

A classic example of a loafing study is one conducted by Latané, Williams, and Harkins (1979). Participants were told to cheer and clap as loudly as possible. They did this alone, or in two-person or six-person "pseudogroups." They were called pseudogroups because participants wore blindfolds and headsets transmitting continuous white noise, so they were merely conscious of the fact that others were present cheering and clapping as loudly as possible, but they were not trying to coordinate their own cheering and clapping with that of others. There was a reduction in effort for participants in pseudogroups—29% and 60% in the case of two- and six-person groups respectively.

Because loafing reflects reduced motivation, anything that increases motivation to work hard can reduce loafing. Research finds that personal identifiability by the experimenter, personal involvement in the task, partner effort, intergroup comparison, and a highly meaningful task in association with expectation of poor performance by coworkers all reduce loafing. Sometimes people may even work harder in a group than alone, specifically to compensate for their perception that others will loaf on important tasks or in important groups. This produces a compensation effect—increased effort on a collective task to compensate for other group members' actual, perceived, or anticipated lack of effort or ability (Williams & Karau, 1991). Overall, loafing is reduced, and often reversed, when a task is meaningful and has a subjectively important end product and, perhaps most critically, when people identify strongly with a personally significant group in which they are working (Fielding & Hogg, 2000).

DECISION MAKING

Many of the decisions that affect us most are made by groups, ranging from organizational committees, to government or global decision making councils. It is difficult to envisage social life without participating in or being influenced by group decision making.

Decision Rules

Decision-making groups typically bring together people who have different opinions, and there is discussion and exchange of information that generates a group decision. According to Davis's (1973; see Miller, 1985) model of social decision schemes, five context- and task-specific rules, which can be formal but often are implicit, determine the process and criteria for arriving at a group decision.

Unanimity is a painstaking process where all members must be in agreement—this is the explicit rule that juries use to deliver a verdict in criminal cases in the U.S. legal system in almost every state. *Majority wins* is a quicker process where the group's decision is whatever the majority opinion is. *Truth wins* is a process where the group determines the objectively correct position and adopts that as its decision. *Two-thirds majority* specifies that a group decision is reached when two-thirds of members are in agreement. *First shift* allows a group decision to be reached after deadlock; the first opinion shift that a group member makes locks in the group decision.

Although employing a strict rule, such as unanimity, requires a significant amount of time and work, group members tend to be more satisfied with the decision and with their group, and like fellow members more, after they reach a decision than in a simple majority-rules situation. However, because unanimity requires that everyone has to agree, it also can lead to extreme decision making (Miller, 1985). The pressure to reach a unanimous decision paired with extremist voices can lead to group polarization, reflected in the extremity of the decision that the group reaches.

Polarized and Extreme Group Decisions

Groups typically perform an averaging function that leads to a behavioral and attitudinal convergence on the group mean (Sherif, 1936). However, groups sometimes can polarize to be more extreme than their individual members. Group polarization is the tendency for groups to reach more extreme decisions and adopt more extreme behaviors in the direction favored by the group than would be expected from the mean of the initial individual positions of the group's members (e.g., Isenberg, 1986).

There are three explanations for group polarization: (1) social comparison and cultural values, (2) persuasive arguments theory, and (3) self-categorization theory. The social comparison and cultural values perspectives (e.g., Sanders & Baron, 1977) assume that people seek social approval and try to avoid social censure and that fellow group members are a reliable source of information about what is socially desirable. Group discussion reveals which attitude pole (extremely positive or negative) is socially desirable or culturally valued, and group members compete to appear to be strong advocates of that pole—there is a shift in the direction of the group to gain approval and avoid disapproval.

Persuasive arguments theory (Burnstein & Vinokur, 1977) argues that people are persuaded by novel information if it supports their preexisting beliefs. Group discussion exposes individuals to the views and arguments of other group members, some of which they may not have encountered before (novel information). If group members already lean in one direction, then novel arguments will strengthen and entrench their leaning to make each member more extreme, thus shifting the group as a whole to endorse a more extreme and polarized position.

Persuasive arguments theory assumes that group members actively share special knowledge and novel information with the group. However, this may not always happen. Research on "hidden profiles" (Stasser & Titus, 2003) suggests that group members do not always feel comfortable sharing novel and diverse information with the group. The most common paradigm for this research involves providing student participants in simulated decision-making groups with carefully constructed information on the nature of and views held by (hypothetical) other group members and then measuring what information the participant is prepared to share with the rest of the group.

Group discussion tends to revolve around information that everyone in the group already knows and is in agreement on. A member with novel information may incur social costs (such as marginalization and censure) by sharing such information with the group because it can pose a threat to preexisting norms and group uniformity. When low-status members share novel information, it is often dismissed and not remembered; such information shared by high-status members or experts tends to be considered and remembered. This finding suggests that effective leadership in decision-making groups often may be needed in order to encourage members to consider all novel information irrespective of its source.

The third explanation of group polarization is provided by self-categorization theory (Turner, Wetherell, & Hogg, 1989). People construct their group's normative position (the in-group prototype) on an issue by identifying not only similarities among members but also key differences between their group's position and that of an out-group that they strive to be distinct from. This process of metacontrast produces an in-group position that is more extreme than the in-group mean position in a direction that is polarized away from the relevant out-group. The cognitively represented position of the in-group is thus polarized, and the process of self-categorization associated with group identification ensures that each member internalizes this position as their own and expresses it through their behavior. The group shifts in line with the polarized norm. For example, members of a committee who identify strongly as liberals would adopt and voice a more extreme liberal group position on, say, immigration when the salient out-group is “conservatives” and a more conservative position on immigration when the salient out-group is “socialists.”

Groupthink

Unanimous decision-making groups can, as mentioned, be very satisfying to participate in, as they create a sense of cohesion and common purpose. However, there is one key pitfall of such groups—they can produce groupthink (Janis, 1982). Groupthink is a mode of thinking in which the desire to reach unanimous agreement overrides the motivation to adopt proper rational decision-making procedures. Groupthink tends to occur in highly cohesive groups that have strong directive leadership and are under pressure to make a very consequential decision quickly.

In a rush to come to consensus under threat from external pressures, group members may be especially likely to ignore dissenting opinion in favor of uncritically accepting in-group norms set by their group leaders that support their existing beliefs and wider ideology. They may isolate themselves from the outside world and ignore and fail to seek or accept information from outside sources. Such groups tend to be led by a strong leader with whom members do not wish to disagree. In the rare case of dissent, such groups exert pressure toward consensus, avoid exposure to outside opinions, and stifle creative thought from within.

Groupthink produces an overwhelming desire among group members to reach unanimous agreement that leads to disregard for rational decision-making procedures and thus produces poor decisions that result in unfavorable outcomes that can have widespread catastrophic consequences. Janis (1982) originally compared the 1961 Bay of Pigs fiasco with the 1962 Cuban Missile Crisis—the former was a clear example of groupthink; the latter was an example of optimal group decision making. More recently, it has been argued that groupthink played a role in the lead-up to the 2003 invasion of Iraq, a decision famously based on the U.S. government’s “decision” that Iraq did indeed have weapons of mass destruction (McQueen, 2005).

Groupthink and risky decision making can be avoided by appointing a devil’s advocate to highlight contrasting points of view. In addition, leaders can encourage opposition to be voiced within the group and encourage the group to consider opinions and viewpoints originating outside the group.

COMPOSITION, STRUCTURE, AND INFLUENCE

Groups are internally differentiated—people perform different roles, have different opinions, vary in whether they are considered typical or ideal members, vary in whether they are newcomers or old-timers, and differ in the extent to which they have influence over group members and normative properties of the group.

Group Socialization

One key feature of groups is a process of socialization that transforms prospective and new members into full members and a mirror process in which members gradually disconnect from the group to become ex-members. According to Levine and Moreland (1994; Moreland & Levine, 1982), group socialization is a dynamic and reciprocal influence process of evaluation and commitment between individual and group. A prospective member and the group itself evaluate the relative costs and benefits of membership; if both parties consider the benefits to outweigh the costs, then there is a role transition (often associated with rites of passage) from prospective member to new member, which is associated with enhanced mutual commitment. The group accommodates the new member, and the new member conforms to the group's norms.

The same process is associated with other role transitions, for example, from newcomer, to full member, to old-timer, to ex-member. Assuming that the needs of the group and the member are met and there is commitment between the group and member, maintenance occurs whereby the member's role within the group is established. This model allows for the divergence of the member from the group and exit or reentry to the group. Interestingly, new members often are afforded more leeway in the group than are old-timers. This is because established group members are expected to uphold group norms (Levine & Moreland, 2002); if they violate group norms, the group judges them more harshly than it judges new members performing the same acts (Pinto, Marques, Levine, & Abrams, 2010).

Group Norms and Deviant Individuals

Norms map out the contours of groups and prescribe membership-appropriate thoughts, feelings, and actions (Hogg, 2010). The more strongly people identify with their group, the more they adhere to the group's norms, cognitively represented as a prototype, and the more they enforce the group's norms on fellow members (Abrams & Hogg, 1990; Turner, 1991). Because group prototypes are the defining features of our groups, we strive for them to distinguish clearly between in-group and out-group, and we are highly attentive to how we ourselves and others measure up to the prototype—we are attentive to prototypicality.

Distinctive prototypes satisfy a key group and identity motivation: to reduce uncertainty about the group and thus about ourselves as group members (Hogg, 2007). One consequence of this identity-uncertainty reduction motivation is that in-group members who deviate from group norms in a direction that places them effectively on the boundary between in-group and out-group threaten the normative clarity of the group and are reacted to harshly, more harshly than members who deviate in a direction away from the out-group (Marques, Abrams, & Serôdio, 2001). Norm violators pose a fundamental threat to the integrity and cohesion of a group; thus, central and prototypical group members tend to work to enforce group members' normative behavior and punish deviants through marginalization and ostracism (Abrams, Marques, Bown, & Henson, 2000).

People do not only like their in-group prototype to be clear and distinctive; they also like it to reflect evaluatively positively on the group because this satisfies a fundamental human motivation to feel good about one's group and thus one's identity (e.g., Abrams & Hogg, 1988). One consequence of this social identity enhancement motivation is that marginal group members who have undesirable human traits and thus are disliked are treated more harshly if they are in-group members than out-group members (they are "black sheep"). The opposite is true for marginal members who have desirable human traits and are likable (Marques & Páez, 1994).

Conformity and Obedience

Why and how do people conform to group norms? One obvious reason is because a powerful individual in a position of authority orders them to behave in a certain way. In this situation, people simply obey commands rather than conform to norms. Milgram's classic program of studies conducted in the early 1960s showed that ordinary people were quite capable of obeying an authority figure who ordered them to administer deadly electric shocks to an unwitting subject who simply made mistakes in a word association task (Milgram, 1963; cf. Burger, 2009).

Subsequent research has identified a range of variables that influence blind destructive obedience: Obedience is strongest if the authority is legitimate and immediate (physically close by), if the victim of the destructive behavior is remote (effectively out of sight and out of mind), and if the orders gradually build from less to more destructive behavior (the psychology of sunk costs) (Milgram, 1974; see Martin & Hewstone, 2007). Of most relevance to this chapter, however, is the finding that obedience is absolute if people learn that others similar to them are also obeying, but it drops to practically zero if they learn that others are disobedient—behavioral group norms are a powerful influence on individual behavior.

Obedience is not conformity. Conformity to group norms occurs through at least three different influence processes (Abrams, Wetherell, Cochrane, Hogg, & Turner, 1990; Hogg, 2010): (1) normative influence, (2) informational influence, and (3) referent informational influence. Normative influence is a process where people publicly go along with the group to obtain social approval and avoid disapproval; informational influence is a process where people turn to others as a reliable source of information about the nature of reality (Deutsch & Gerard, 1955). Referent informational influence is a process where people turn to others as a reliable source of information about the self-relevant social identity–defining group norm in a specific context and then internalize that norm to determine their own perceptions, attitudes, and behaviors (Abrams & Hogg, 1990; Turner et al., 1987).

Minority Influence and Social Change

Influence in groups tends to be a process whereby group norms and identity embody the attributes of the majority of group members and members are influenced to conform to the majority. This finding suggests that minority or countermajority viewpoints and identities gradually wither away, and there is little prospect for social change at the small group or societal level. The social identity analysis of groups, norms, and influence belies this process because, as we have seen above, it ties identities and norms to salient intergroup comparisons—groups change their norms and identity as the social comparative context changes.

A more complete analysis of social change is provided by minority influence research (e.g., Martin & Hewstone, 2008; Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994). The main premise is that a numerical or power minority that promulgates a consistent, not rigid, message that is not transparently self-interested can change the views of the majority and thus create dramatic social change (Moscovici, 1976; Mugny, 1982). The process is one of conversion—the minority message is publicly rejected, but because it is novel and consistent, it privately attracts attention and cognitive resources that ultimately lead to a sudden conversion in favor of the minority viewpoint (Moscovici, 1980). Research suggests that a minority that can be viewed by the majority as an in-group minority is more effective because the majority extends it some leniency, and this process at least lowers some cognitive barriers to consideration of its position (Crano & Seyranian, 2009).

LEADERSHIP

Although people in groups ultimately are influenced by group norms, it is real people (as opposed to abstractions) who do the actual influencing and to whom we turn for guidance about how we should behave as group members. The study of influence is inextricable from the study of leadership (Hogg, 2010). If we assume that an effective leader is someone who has disproportionate influence in a group in motivating followers to adopt his or her vision for the group and exert effort on behalf of that vision, then we know a great deal about the psychology and social psychology of leadership (Yukl, 2010).

Effective leaders typically score high on the key personality constellations of conscientiousness, extraversion/surgency, and intellect/openness to experience and are considered charismatic, innovative, and transformational (Judge, Bono, Ilies, & Gerhardt, 2002). They typically are good at what they do for the group, have high status in society, and match our schema of what attributes a leader should have (Lord & Hall, 2003; Ridgeway, 2003). Thus, people belonging to social categories that societal stereotypes deem as not well suited to leadership—for example, women—can have an uphill struggle to lead (Eagly & Karau, 2002). Effective leaders are able to manage relationships within the group (relationship orientation), focus members on the group task (task orientation), and know in what situations to prioritize one or other or both of these orientations (Fiedler, 1964). Such leaders are proficient at managing their transactions with followers such that followers feel rewarded, trusted, and valued (Graen & Uhl-Bien, 1995; Tyler & Lind, 1992); and in return, followers reward their leader with the trappings of leadership and the ability to lead and be innovative (Abrams, Randsley de Moura, Marques, & Hutchison, 2008).

Leaders are not mere managers of group activities and tasks; they also define the group's identity—they serve a social identity–defining function that is critical in groups that are central to a person's social identity (Hogg & Van Knippenberg, 2003; Hogg, Van Knippenberg, & Rast, 2012b). Under these circumstances, one of the most critical influences on leadership effectiveness is the extent to which the group perceives the leader to be a prototypically central group member (one of us) (Barreto & Hogg, 2017). Members trust prototypical leaders to have the group's best interest at heart; thus, prototypical leaders are allowed to be normatively innovative, and members turn to them to define the group's identity. This social identity dimension of leadership is also critical when, as is often the case, leaders have to provide leadership across a deep divide between very different groups (e.g., uniting Sunni, Shia, and Kurds in Iraq). The challenge of intergroup leadership is to construct a collaborative relationship and common bonds that do not threaten each group's cherished identity distinctiveness within the larger collective (Hogg, 2015a; Hogg, Van Knippenberg, & Rast, 2012a).

INTERGROUP RELATIONS

Thus far we have focused primarily on intragroup processes—what happens among people within groups. Now we change focus to intergroup relations—the way in which people who belong to social groups or categories perceive, think about, feel about, and act toward and interact with people in other groups (Hogg, 2013; Yzerbyt & Demoulin, 2010). However, it is important to recognize that, in reality, what happens between groups affects what happens within groups and vice versa. Inter- and intragroup processes are largely inextricable, so the distinction is chiefly a matter of focus or emphasis.

However, there is one preeminent cluster of features of intergroup relations that differentiates it from intragroup phenomena. Intergroup relations are intrinsically competitive, ethnocentric, and discriminatory (Dovidio & Gaertner, 2010). Groups compete with one another over resources and prestige; they feel they are superior to one another and have negative stereotypes of one another; they behave to varying degrees in ways that discriminate against or stigmatize out-groups; and they protect and differentiate their identity relative to salient outgroups.

Even merely being categorized as a member of a group can, as we show, generate the foundations of these symptoms (Tajfel, 1970), and some research has suggested that this association between being in a group and discriminatory intergroup perception and behavior is an automatic one that is hard-wired in the brain (Otten & Wentura, 1999). This fundamental intergroup orientation can translate into more extreme and harmful forms when groups feel their resources, prestige, or identity are threatened (Mummendey & Otten, 1998). How groups behave in these situations is shaped by the realities of their relations to other groups (Ellemers, 1993; Hogg & Abrams, 1988; Tajfel & Turner, 1986).

Intergroup relations capture many of society's greatest ills, including prejudice, discrimination, stigma, dehumanization, genocide, and war. Not surprisingly, this topic is not only of great scientific interest but also of enormous social relevance.

PERSONALITY AND INDIVIDUAL DIFFERENCES

Given the suffering caused by some extreme forms of intergroup behavior, specifically prejudice and its behavioral consequences, finding out why people are prejudiced is a large part of intergroup research. Perhaps only some of us with particular personality configurations are capable of behaving in a prejudiced and discriminatory way. The search for a "prejudice personality type" has preoccupied social psychologists and produced a number of accounts of personality and individual differences. These explanations include the authoritarian personality theory and social dominance theory.

Authoritarianism and Closed-Mindedness

Foremost among personality explanations is the authoritarian personality theory of Adorno, Frenkel-Brunswik, Levinson, and Sanford (1950). Adorno and colleagues adopted a psychodynamic framework to argue that early childhood rearing practices that are harsh, disciplinarian, and emotionally manipulative produce people who are obsessed by status and authority, intolerant of ambiguity and uncertainty, and hostile and aggressive toward weaker others. These people have an authoritarian personality that predisposes them to extreme forms of intergroup behavior.

Research on the authoritarian personality confirms the existence of such a syndrome but does not provide good evidence for its origins in early child rearing or for its relationship to prejudice and discrimination.

(See Duckitt, 2000.) People who do not have an authoritarian personality can be prejudiced, and people who do have an authoritarian personality can be free of prejudice. For example, Pettigrew (1958) conducted a now-classic study in which he administered a survey measuring authoritarianism and intergroup attitudes to white South Africans and to white Americans from the North and the South of the United States. He found that prejudice was less related to personality than it was to socialization within a culture of prejudice that legitimized prejudice as the background to everyday life.

Although Pettigrew's social contextual perspective is now widely accepted by social psychologists who study prejudice and intergroup relations (e.g., Reynolds, Turner, Haslam, & Ryan, 2001), other explanations that are oriented largely or somewhat to personality and individual differences do exist; however, these explanations are divorced from the psychodynamic grounding of the authoritarian personality syndrome. For example, Altemeyer's (1998; also see Duckitt, Wagner, du Plessis, & Birum, 2002) theory of right-wing authoritarianism characterizes authoritarianism as an ideology that maintains the status quo by promoting conventionalism (adherence to societal conventions that are endorsed by established authorities), authoritarian aggression (support for aggression toward social minorities and deviants), and authoritarian submission (submission to society's established authorities).

In addition, a number of perspectives argue that some people are closed-minded, intolerant of uncertainty, or in need of cognitive closure, and this inclines them toward ethnocentrism, intolerance, and bigotry (e.g., Kruglanski, Pierro, Mannetti, & De Grada, 2006; Kruglanski & Webster, 1996; Rokeach, 1948).

Social Dominance and System Justification

Others, from a social dominance theory perspective, focus on power-based intergroup relations and the extent to which people accept or reject societal ideologies or myths that legitimize hierarchy and discrimination or that legitimize equality and fairness (e.g., Sidanius & Pratto, 1999). People who strongly endorse social hierarchies grounded in group inequalities are high in social dominance orientation. These kinds of people are more inclined to be prejudiced than are people who have a low social dominance orientation.

System justification theory (e.g., Jost & Hunyadi, 2002; Jost & van der Toorn, 2012) also considers individual support for social hierarchies, specifically examining the social conditions that cause people to resist social change and instead justify and protect the existing social order, even if it maintains their own group's position of disadvantage.

COMPETITION AND COOPERATION BETWEEN GROUPS

An alternative view on intergroup relations is that the tenor of intergroup perceptions and behavior cannot be deduced from personality and individual differences but reflects the nature of the relationship that groups believe they have with one another. Muzafer Sherif (1962) famously argued that "we cannot extrapolate from the properties of individuals to the characteristics of group situations" (p. 5).

Sherif's (1958) realistic conflict or interdependence theory of intergroup relations argues that group-based goals and people's perceptions of their relationship to one another with respect to goal attainment drives intergroup behavior. If two groups have the same, mutually exclusive goal (there is a zero-sum goal relationship), intergroup relations will be competitive and hostile. If two groups share a common superordinate goal that they can achieve only through cooperation, intergroup relations will be harmonious.

Sherif and his colleagues (e.g., Sherif, Harvey, White, Hood, & Sherif, 1961; see Sherif, 1966) tested interdependence theory in their classic Robbers Cave field experiments. They manipulated goal relations between rival groups that they created at boys' camps in the United States. These experiments provided support for realistic conflict and interdependence, even demonstrating the ability to overcome intergroup hostility through shared superordinate goals. Subsequent research has robustly supported the idea that intergroup goal relations have a significant impact on the nature of intergroup perceptions and behavior (e.g., Blake & Mouton, 1961; Brewer & Campbell, 1976; Fisher, 1990).

SOCIAL CATEGORIES AND COGNITIVE PROCESSES

Social Categorization

Intergroup goals and goal relations are important, but where there are such goals, there must be groups, and perhaps the very fact of being in a group, even in the absence of competitive goal relations, is sufficient to generate characteristic intergroup perceptions and behavior. It turns out that this is true.

The minimal groups experiments investigated how mere categorization into different groups can cause group members to favor their own group over other groups (Tajfel, 1970; Tajfel, Billig, Bundy, & Flament, 1971). Participants who were simply categorized into groups on a random or minimal basis favored their own group over the out-group and indicated a sense of ethnocentrism and belonging to their minimal in-group—this despite the fact that membership was anonymous and the groups had no history or future and were not in competition over resources (cf. Abrams & Hogg, 2010). As mentioned earlier, this minimal group effect may be an automatic consequence of social categorization, and it becomes more extreme under conditions of intergroup conflict and threat (e.g., Mummendey & Otten, 1998; Otten & Wentura, 1999).

The minimal group studies suggested that the act of viewing oneself as a group member may impact one's sense of identity and that this may play a key role in group processes and intergroup relations. This idea prompted the development of social identity theory, which we describe in the section titled "Self and Social Identity."

Automatic Schema Activation

People cognitively represent social categories as schemas that describe the attributes of the category and the relationships among those attributes. Schemas vary from concrete exemplars of the category, to abstract fuzzy sets of loosely related attributes (prototypes) (e.g., Rosch, 1978). Categories themselves vary in entitativity (the degree to which they resemble a tightly organized, distinctive, and cohesive, unitary construct; Hamilton & Sherman, 1996). Entitativity is associated with a range of group processes and intergroup behaviors, such as conformity to group norms, rejection of deviants, hierarchical leadership, and polarized intergroup perceptions and behavior. (See Yzerbyt, Judd, & Corneille, 2004.)

Perceptual cues, in particular distinctive visual cues, prompt us to categorize people and imbue them with the attributes described by our schema of that group (Macrae & Quadflieg, 2010). The entire process can be slow and deliberate, but in general it is fast and automatic (Fiske & Neuberg, 1990). Stereotyping of out-group members may be largely an automatic categorization-contingent process that we have little control over (Bargh, 1994). Category attributes are implicitly linked to and associated with categories (Greenwald et al., 2002), and this process largely occurs in that part of the brain (the amygdala) where automatic cognitive processing in general occurs (Lieberman, 2007).

This automatic process may be moderated by a number of factors. For example, the category-attribute link is weaker among those who are less prejudiced (Lepore & Brown, 1997) and those who are able to take the perspective of the target (Galinsky & Moskowitz, 2000). There is also evidence for a stereotype rebound effect—if people consciously think about the automatic category-stereotype link, there is a paradoxical strengthening of the link that increases automatic stereotype activation (e.g., Zhang & Hunt, 2008).

Accentuation and Illusory Correlation

Social categorization also causes us to perceptually accentuate similarities among members of the same category and differences between members of different categories on dimensions that distinguish the categories (Tajfel, 1959). This accentuation effect is a general consequence of categorization, but one that is asymmetrical because we generally view out-groups as more homogenous than in-groups (e.g., Judd & Park, 1988).

One reason for the asymmetry is that we are more familiar with the in-group and therefore have more individuating information about in-group members than out-group members (Linville, Fischer, & Salovey, 1989). However, it may also reflect strategic motives (Jones, Wood, & Quattrone, 1981). For example, active minorities often consider themselves to be more homogenous than the majority out-group, a view that clearly is functional as active minorities need to be consistent and consensual to survive and have a realistic opportunity to create social change (Mugny, 1982; Simon & Brown, 1987).

Related to the accentuation effect is the illusory correlation effect (Hamilton & Sherman, 1989) where people associate distinctive behaviors with distinctive categories, thus laying a foundation for correlating unfavorable attributes with minority groups. (Both unfavorable attributes and minority groups are believed to be perceptually distinctive.)

SELF AND SOCIAL IDENTITY

Groups, as social categories, define who we are—our sense of identity and self in society. Thus, processes associated with self-conception play a fundamental role in group processes and intergroup relations. This general idea is the foundation of social identity theory, which is a general theory of group processes and intergroup behavior (Tajfel & Turner, 1986; Turner et al., 1987; also see Abrams & Hogg, 2010; Hogg, 2006, 2012a; Hogg & Abrams, 1988).

People cognitively represent social categories as prototypes—fuzzy sets of attributes that capture intragroup similarities and accentuate intergroup differences, on dimensions that both accentuate in-group entitativity and secure a positive social identity for the in-groups relative to relevant out-groups. The process of social categorization depersonalizes perception of self and others, which means that rather than viewing ourselves and others as idiosyncratic individuals, we view ourselves and others in terms of the relevant in-group or out-group social identity and associated prototype. In the case of self-categorization, one's self-concept is transformed to embody the prototypical attributes of the contextually salient in-group social identity. This means that our attitudes, feelings, and behavior conform to and express the in-group norm. Because the in-group prototype is key to our identity in group contexts, people in groups are highly attentive to reliable information about the prototype and about how well we and others match the prototype.

Because social identities not only define but evaluate who we are, people strive to establish or maintain the evaluative superiority of their own group over relevant other groups—there is a fierce intergroup struggle for evaluatively positive group distinctiveness. This struggle is, however, tempered by people's understanding of the nature of the relations between their group and relevant out-groups (cf. Ellemers, 1993). In particular, people attend to status differences and the stability and legitimacy of such differences, to the permeability of intergroup boundaries and thus the possibility of passing psychologically from one group to the other, and to the existence of achievable alternatives to the status quo.

For social identity theory, group behaviors (conformity, stereotyping, ethnocentrism, in-group favoritism, intergroup discrimination, in-group cohesion, etc.), as distinct from interpersonal behaviors, occur when social identity is the psychologically salient basis of self-conceptualization; and the content of group behavior rests on the specific social identity that is salient. Social identity is context specific insofar as different social identities are salient in different social contexts.

Social identity–grounded group phenomena are motivated by at least three basic motivational processes: (1) to reduce feelings of uncertainty about one's self and identity through being affirmed as a central member of a highly entitative group (uncertainty-identity theory; Hogg, 2007); (2) to elevate one's self-esteem by identify strongly with a (high-status) group that provides a positive social identity and sense of self (self-esteem hypothesis; e.g., Abrams & Hogg, 1988); and (3) to achieve an optimal balance between contrasting drives to be distinctive and to be assimilated within the group (optimal distinctiveness theory; e.g., Leonardelli, Pickett, & Brewer, 2010).

Social identity theory has made an enormous impact on the social psychology of intergroup relations and also has contributed significantly to a revival of research on group processes in general (Moreland, Hogg, & Hains, 1994; Randsley de Moura, Leader, Pelletier, & Abrams, 2008).

GROUP-BASED AFFECT AND EMOTION

Social psychological analyses of intergroup relations tend to focus on social cognitive processes. However, intergroup relations are notable for the strong feelings and associated emotions people have about their own group and about out-groups. The most troublesome aspects of intergroup relations are precisely those that involve strong emotions and powerful affect.

Intergroup emotions theory draws on social identity and self-categorization processes to show how people who identify with the same group can collectively experience very similar emotions (e.g., Mackie, Maitner, & Smith, 2009; Smith, Segar, & Mackie, 2007). People in salient groups feel other people's emotions as their own because self-categorization merges self and other via prototype-based depersonalization—people experience or include the group and its members as part of the self (e.g., Swann, Jetten, Gomez, Whitehouse, & Bastian, 2012;

Tropp & Wright, 2001). In this way, intergroup feelings (which are often negative) can readily become powerfully consensual collective intergroup feelings. Similarly, positive in-group feelings can transform into consensual positive regard and in-group solidarity.

The aversive emotions of guilt and shame, when experienced collectively on behalf of one's group as a reaction to the way it has treated another group, impact the extent to which a group intends to perform acts of reparation, specifically public apologies and political action intentions. For example, Doosje, Branscombe, Spears, and Manstead (2006) examined how a dominant group experiences collective guilt (a feeling that "we" have unjustly disadvantaged and caused suffering to "them"). If members identify strongly with the group, the identity is central to self-conception, and members believe that the group's position of superiority is illegitimate because it rests on the group's violation of a moral value to which the group adheres.

STEREOTYPING AND PREJUDICE

When we think of the attitudes people have of out-groups, we think of stereotypes. Stereotypes and prejudices are generalized images that members of a group share about an out-group (Macrae, Stangor, & Hewstone, 1996; Oakes, Haslam, & Turner, 1994). Although we typically think of them solely as unfavorable attitudes about an out-group and its members, they can also be positive and about the in-group. Using the language of social identity, stereotypes are effectively shared group prototypes. As such, the process of stereotyping is, as described, the assignment of group prototypical characteristics to an individual on the basis of that individual's category membership alone.

Although stereotyping is largely automatic (Dijksterhis, 2010), it can also be deliberative, because stereotypes serve strategic social functions (Oakes, Haslam, & Turner 1994). For example, they may emerge to justify actions that have been committed or planned by one group against another group (if one group exploits another, it may be useful to justify this action by developing a stereotype of the out-group as unsophisticated and dependent), and stereotypes become more extreme and more resistant to change under conditions of intergroup conflict.

Where the expression of specific stereotypes and prejudices is socially or legally proscribed, people are adept at hiding their deep-seated inner attitudes. This can make it difficult to know or measure people's prejudices. However, social psychologists have developed an array of unobtrusive and implicit measures of prejudice. The most well-known and widely used is the Implicit Association Test, in which implicit cognitive associations are elicited by measuring how quickly people press computer keys to associate evaluative words with social categories (Greenwald, McGhee, & Schwartz, 1998; also see Chapter 12 in this volume).

Modern Forms of Prejudice

Where the social context encourages or does not inhibit the expression of prejudices, people feel quite comfortable being overtly prejudiced in what they say and what they do. In many parts of the world and sectors of society, the "old-fashioned" public expression of prejudices is common and unremarkable. However, social norms relating to prejudice can change relatively quickly while the deep-seated cognitive associations between social categories and stereotypical attributes laid down in childhood do not. This creates a situation where people realize that prejudice is unacceptable but nonetheless they harbor prejudiced attitudes.

This latter situation is common in modern Western societies that promote an ethos of tolerance and egalitarianism enshrined in social norms and legislation. Here, many people conceal their deep-seated negative intergroup attitudes and instead express prejudice in more modern forms (e.g., Dovidio & Gaertner, 1998). For example, a long history of racism produces deep-seated racial prejudice, fears, and suspicions. However, a tradition of tolerance and egalitarianism has become enshrined in social norms and legislation that proscribe racist behavior. For many people, therefore, there is an uncomfortable psychological conflict between two sets of contrasting beliefs. People can resolve the conflict by avoiding the racial outgroup, avoiding the issue of race, denying the existence of disadvantage, opposing preferential treatment, among other methods. In the context of gender, old-fashioned hostile misogyny can be transformed into modern benevolent sexism that characterizes women as warm and caring but not as competent as men (e.g., Glick & Fiske, 1996; Swim, Aikin, Hall, & Hunter, 1995).

Explanatory Functions of Stereotypes

Specific stereotypes rarely stand in isolation; they are embedded and networked within a wider social explanatory framework, an ideology or worldview (cf. Duckitt et al., 2002). For example, Crandall (1994) has shown how unfavorable attitudes toward people who are obese are configured by a wider ideology about "fat" people as a social category of people who "choose" not to take responsibility for their health or appearance. Negative attitudes toward categories such as the unemployed, and even rape victims, may be configured by a wider ideological belief in a just world (cf. Furnham, 2003), which ultimately perpetuates stigmatization based on victim blaming.

This explanatory dimension of stereotypes probably rests on a wider and fundamental human drive to attribute causes, however inaccurate, to people's behavior—causal explanations of our social world allow us to anticipate and plan action (e.g., Hilton, 2007). However, such attributions tend to be impacted systematically by social categorization and group membership motives, such that we are prey to an ethnocentric attribution bias called the ultimate attribution error (Islam & Hewstone, 1993; Pettigrew, 1979). Good acts are dispositionally attributed when performed by an in-group member and situationally attributed when performed by an out-group member, and vice versa for bad acts.

DISCRIMINATION, STIGMA, AND DISADVANTAGE

Intergroup relations are not only a matter of ethnocentric and in-group–favoring perceptions, attitudes, and social explanations but also can involve behavior that materially disadvantages and stigmatizes out-groups and involves intergroup aggression (Dovidio, Glick, Hewstone, & Esses, 2010).

From Prejudice to Destructive Intergroup Behavior

A key question concerns the conditions under which groups express prejudiced attitudes as behavioral discrimination and, relatedly, the extremity and form of the behavior. We have seen that prejudices can be concealed and that it can be difficult to detect them (cf. Implicit Association Test; Greenwald, McGhee, & Schwartz, 1998) and that where progressive societal norms prevail, prejudice can be expressed in modern and "benevolent" forms that do not readily look like old-fashioned prejudice (Glick & Fiske, 1996). We also have seen that in-group–favoring behavior can be elicited by the mere existence of even minimal social categories (Tajfel, 1970), that in-group favoritism morphs into out-group discrimination under conditions of intergroup conflict and identity threat (e.g., Mummendey & Otten, 1998), and that competitive zero-sum intergroup goals accentuate conflict and destructive intergroup behavior (Sherif, 1966).

Social identity theory offers a textured analysis, which we briefly alluded to earlier, of how groups strive to protect and enhance their status and prestige and thus the evaluative favorability of social identity (Tajfel & Turner, 1986; see Ellemers, 1993; Hogg & Abrams, 1988). People have beliefs about the stability and legitimacy of their group's relative status in society, about the permeability of group boundaries that might let them pass out of one group and be admitted into a more prestigious group, and about the realistic probability of achieving an alternative more favorable status quo. Direct conflict between groups is most likely where perceived permeability is low and where the status quo is considered illegitimate and unstable and where an alternative social order and associated strategies for achieving it can be envisaged.

Extreme intergroup behavior commonly arises when people feel that a social identity that is central to their self-concept is unraveling and becoming fuzzy and unclear; this creates an unsettling sense of uncertainty about one's identity and one's very sense of self. People react strongly to reduce identity uncertainty (Hogg, 2007) or more general personal uncertainty (Van den Bos, 2009) and to establish some sense of cognitive closure (Kruglanski & Webster, 1996). This resolution of uncertainty and uncertainty-related motivations, taken to extremes, can generate a group-centric constellation of behaviors, such as zealotry, xenophobia, extreme ethnocentrism, ideological orthodoxy, support for autocratic leadership, and intolerance of out-groups and nonconformist in-group deviants (e.g., Hogg, 2014; Kruglanski et al., 2006).

Dehumanization and Intergroup Aggression

Intergroup behavior is most destructive to individuals and society when it manifests as aggression and ultimately genocide (Staub, 2010). What enables people to go so far as to engage in or support mass intergroup atrocities?

According to Haslam, Rothschild, and Ernst (1998), the key factor is essentialism and dehumanization—a process where people attribute stereotypical attributes to immutable underlying properties (often viewed as genetically "fixed") of a certain group. When this process of essentialism focuses, as it often does, on highly negative out-group attributes that effectively diminish the targets' humanness, a process of dehumanization has occurred (Haslam, 2006; Haslam, Loughnan, & Kashima, 2008). Viewing people as nonhuman makes it easier both to psychologically and physically harm them and to commit widespread intergroup atrocities, such as genocide.

Stigma and Disadvantage

Intergroup relations tend to position groups in a status hierarchy, placing some groups higher than others in the hierarchy. Of course, lower-status groups can contest their status position, but often doing so is difficult or even dangerous because the dominant group has a monopoly of power in society to define social reality and, in more extreme cases, to aggressively enforce it. One consequence of this status dynamic is that subordinate groups can experience an enduring sense of disadvantage and stigma in society.

Stigmatized individuals are aware that others may judge and treat them stereotypically, and thus, on important tasks, they worry that they may confirm others' negative stereotypes of their group. This anxiety associated with stereotype threat (Steele, Spencer, & Aronson, 2002; also see Inzlicht & Schmader, 2011) often impedes and harms performance on stereotype-defining tasks and thus does confirm others' negative stereotypes.

However, stigmatized groups and their members are remarkably resilient. Although they are vulnerable to low self-esteem, diminished life satisfaction, and in some cases depression, most members of stigmatized groups are able to weather the assaults and maintain a positive self-image (Crocker & Major 1989). They can accomplish this by denying that they have personally experienced discrimination and stigmatization (e.g., Major, 1994), or they can reconcile themselves to their plight by believing that it is preferable to accept the status quo than to confront the uncertainties and dangers of struggling for social change (e.g., Hogg, 2007; Jost & Hunyadi, 2002).

COLLECTIVE ACTION AND SOCIAL PROTEST

Although disadvantaged and stigmatized groups have an impressive armory of protective strategies to redirect energy from direct intergroup conflict, doing so is not always effective. When deprivation is acute and a recipe for effective social change is available, disadvantaged groups do challenge the status quo by political means or through social protest or other collective behaviors including demonstrations, riots, and uprisings.

As discussed in the subsection of this chapter titled "Deindividuation," social psychology traditionally has viewed crowd behavior as an irrational expression of primitive human emotions in situations where people are anonymous and unaccountable and thus freed from the constraints of societal norms (LeBon, 1908; Zimbardo, 1969). However, also as described earlier, other more recent perspectives have noted that the apparent volatility of crowd events may be largely because they are contexts in which people are strongly regulated by social identity but are not quite clear about how they as group members should behave in a novel context; they are not sure what the situation specific group norms are (Reicher et al., 1995). Crowds are actually rational goal-oriented gatherings of people who share a social identity; they are better characterized as collective celebration or protest.

This line of thinking poses the question of what drives a group to take collective action as a social protest aimed at social change. An important factor here is a sense of relative deprivation. When a group has rising expectations but experiences a relatively sudden drop in attainments, it feels frustrated and deprived, and this translates into collective protest (Davies, 1969; Gurr, 1970). Groups also rise up when they feel that they, as a group, are deprived relative to other comparable groups in society; they experience fraternalistic (group-based) relative deprivation that translates into collective action to redress the problem (Runciman, 1966; Walker & Pettigrew, 1984). However, even under these circumstances, as proposed by social identity theory and discussed earlier, the group is most likely to rise up when it can envisage an alternative social order and when it has a clear and concrete plan about exactly how to make it happen.

Social protest research examines how individual discontents or grievances become collective action (e.g., Klandermans, 2002; Stürmer & Simon, 2004; Van Zomeren, Postmes, & Spears, 2008). Protest is generally *for* a social good (e.g., equality) or *against* a social ill (e.g., oppression), even if what is good and what is ill is subjective. Because success benefits an entire group (not just participants in the movement) and there is often a significant cost involved in physically protesting, many people remain on the sidelines as sympathizers and "free ride" on their brethren's efforts. Social protest can, however, be fully understood only in its wider intergroup context, where there is a clash of ideas and ideologies between groups and politicized and strategic articulation with other more or less sympathetic organizations.

Another form of collective action and social protest can occur when a subgroup feels that the norms, objectives, and, ultimately, identity of the parent group have changed and that they as a subgroup have no effective voice to address this change. Under these circumstances, a schism can occur and the subgroup can seek substantial independence and autonomy within the superordinate group (Hogg & Wagoner, 2017; Wagoner & Hogg, 2016) or split off entirely from the parent group and thus effectively create a new hostile intergroup relationship between the new entity and its erstwhile superordinate group (Sani, 2005).

SOCIAL HARMONY BETWEEN GROUPS

The study of intergroup relations often seems like an exploration of human misery—prejudice, discrimination, stigma, conflict, and genocide. However, by understanding the dark side of human nature, social psychology also can tell a more uplifting story—a narrative that points to ways in which intergroup relations can be improved and social harmony can be built (Dovidio et al., 2010).

Prejudice and intergroup conflict can be sustained by lack of face-to-face contact between members of different groups. This observation underpins the contact hypothesis, which argues that relations can be improved if people become familiar with and get to know out-group members as whole human beings similar to themselves rather than as faceless category members (Allport, 1954; Amir, 1969). Pleasant intergroup interaction should change enduring intergroup images through a process of generalization associated with the accumulation of favorable out-group information (bookkeeping), through a sudden encounter with counterstereotypic information (conversion), or through the development of a more textured out-group representation (subtyping) (Weber & Crocker, 1983). People who have pleasant contact with an out-group member who is clearly viewed as being representative or stereotypical of the out-group should develop improved attitudes toward the out-group as a whole, and research supports this finding (Wilder, 1984).

This general idea that contact improves intergroup perceptions and relations helped form the scientific grounding for the U.S. policy of desegregating schooling through busing that began in the 1950s (Schofield, 1991). However, contact works to improve intergroup relations only if it is conducted under very specific circumstance. Pettigrew and Tropp (2006) conducted a meta-analysis of 515 contact studies conducted between 1949 and 2000 with 713 samples across 38 nations. Their analysis revealed that cooperation, shared goals, equal status, norms favoring social equality, and the support of local authorities are boundary conditions that are necessary for contact to improve intergroup relations.

Sadly, it is difficult and often impractical to meet these parameters. Prejudice and conflict typically are deeply rooted, and intergroup contact can be threatening and can confirm people's worst fears and stereotypes. Stephan and Stephan's (2000) integrated threat model outlined four types of threat people experience in anticipation of intergroup contact: (1) realistic threat (a threat to the existence of one's group, well-being, political power, etc.); (2) symbolic threat (a threat posed by the outgroup to one's values, beliefs, morals, and norms); (3) intergroup anxiety (anxiety about personal embarrassment and rejection experienced during contact); and (4) negative stereotypes (imagined or anticipated anxiety based on negative stereotypes of an out-group). Together these perceived threats and associated anxiety can produce intergroup avoidance, negative evaluations, and negative emotions (cf. Aberson & Gaffney, 2009).

Some of these practical and psychological hurdles associated with intergroup contact may be overcome through extended contact, where people observe in-group members successfully engaging in cross-group friendships (Wright, Aron, McLaughlin-Volpe, & Ropp, 1997), or imagined contact, where people role-play and imagine themselves having actual contact with an out-group member (Miles & Crisp, 2014). Extended and imagined contact can improve intergroup perceptions quite effectively.

Nevertheless, even pleasant intergroup contact does not readily generalize from specific contact partners to their group as a whole. People can engage in decategorization, where they treat an out-group member as a unique individual, or they can engage in recategorization, where they recategorize themselves and their partner as members of an overarching shared group. ("Them versus us" becomes "we" under a larger umbrella.) Both strategies reduce intergroup discrimination, but by different routes: Recategorization improves out-group attitudes, whereas decategorization worsens in-group attitudes (Gaertner & Dovidio, 2000).

Recategorization has its own problems; it can threaten subgroup distinctiveness and the integrity of associated social identities and thus enhance identity-protective inter(sub)group conflict. Many organizational mergers fail for precisely these reasons (Cartwright & Schoenberg, 2006), and at the societal/national level, assimilationist strategies aimed at forging a single harmonious cultural or national identity out of many cultural groups can backfire. Social harmony may be better served by a multicultural strategy that avoids the distinctiveness threat raised by assimilationism (Hornsey & Hogg, 2000; Verkuyten, 2006).

CONCLUSION

The study of group processes and intergroup relations is a huge subarea of social psychology, so the account we have given here is by necessity only an overview of key aspects. (For more detail, see: Dovidio & Gaertner, 2010; Dovidio et al., 2010; Hogg, 2010; Levine, 2013; Levine & Hogg, 2010; Stangor, 2016; Yzerbyt & Demoulin, 2010.)

Social groups generally consist of two or more people who at the very least categorize themselves as members of the same group and thus share a social identity. However, groups vary substantially in how they are structured—particularly in terms of size, purpose, distinctiveness, degree of interdependence, shared fate, and similarity of members; and also in terms of the degree to which members emphasize shared category attributes or networks of relationships. People join groups or identify with them for a variety of reasons, chief among which are protection, achievement of shared goals, achievement of a clear and distinct identity, self-esteem through belonging to a socially valued group, and an overarching human need simply to belong.

People in groups can find that they perform well-learned tasks better than alone but poorly learned tasks much worse than alone. People also tend to put in less effort on group tasks (they loaf) unless the task is subjectively important and they identify strongly with the group; and if, in addition, they feel they need to compensate for others who are less competent than they are on a group-defining task. Sometimes being in a group makes people feel anonymous and unaccountable; this sense of deindividuation can lead them to behave antisocially or adhere very strongly to the norms of the group that they feel so immersed in.

A key function of groups is decision making—a process that can be very effective because it benefits from sharing of diverse views and expertise. Groups often have more or less formal decision rules relating individual positions to the group's position: Unanimity is a particularly strict rule that slows decision making but often ultimately enhances satisfaction with the group because everyone feels they have been fully heard. However, decision-making groups have a tendency to make polarized decisions when they already lean in one direction and compare themselves with a group that leans the other way, and these groups have a tendency to make very poor decisions when they are highly cohesive, under pressure, and have strong leadership that together create a strong drive simply to reach agreement.

Groups are not homogeneous; members occupy different specific and more general roles and vary in terms of their centrality and prototypicality. Centrality and prototypicality are defined in reference to the group's norms, where such norms largely define the group and regulate the attitudes, behavior, and identity of members. People who deviate from the group's norms tend not to be liked or trusted, especially if they seem to lean toward a salient out-group's norms and thus blur intergroup boundaries.

Norms are powerful; an emerging norm of disobedience can even combat the strong tendency for people to adhere to the orders of close-by authority figures to do destructive things. Group norms are influential because people do not like to stand out and be disliked by others who are in a majority and because people often view the majority's actions as evidence about the nature of reality or about their social identity. However, minorities also can be influential, particularly if they are distinct entities, viewed as part of the in-group, and provide a consistent message. Group influence tends to work through leadership—effective leaders have persuasive attributes that include being charismatic, but they are also transformational and well versed in managing people and in focusing on the group task, and they know when to prioritize one or the other or both. Effective leaders are also group members who are considered to be prototypical of the group and thus are trusted to do no harm and to define the group and its associated member-defining social identity.

What happens in groups cannot be dissociated from what happens between groups; there is a reciprocal dynamic between intra- and intergroup processes and relations. Intergroup relations are intrinsically competitive and ethnocentric. They are not necessarily destructive or hostile, although research focuses more on this negative aspect. Some individuals are more inclined toward destructive intergroup behaviors—particularly those who are closed-minded and have an authoritarian personality and those who have a strong need for closure and endorse hierarchy. However, situational factors probably are a much more important determinant of intergroup conflict—specifically, the presence of a zero-sum mind-set in which only one group can achieve a valued goal at the expense of the other group. In contrast a non-zero-sum mind-set where both groups have to work together to achieve a shared goal can create intergroup harmony, at least for a time.

The reason for noting "at least for a time" is that the very existence of social categories seems to generate intergroup attitudes that are almost impossible to erase because attributes associated with a category (stereotypes) are elicited automatically by category cues, and we engage in attributional dynamics that ground negative attributes of out-groups in immutable (cf. genetic) properties of the category. Even when people publicly paper over their deeply held prejudices, such prejudices still may lurk deep within the psyche. Social categories also may have such a powerful influence over perception and behavior because social categories define who we are; they construct social identities and ultimately our very sense of who we are in the world. Because social identity is associated with very powerful human motivations to belong, to reduce uncertainty about who we are, and to manage how others evaluate us and how we evaluate ourselves, it is not surprising that differential intergroup behavior that can get quite defensive and assertive is so entrenched and difficult to change.

Prejudice can be expressed overtly, but where societal norms proscribe prejudice, it can be well concealed and expressed in ways that look less like prejudice—for example, by avoiding out-groups or expressing out-group attitudes that on the surface seem benevolent. Prejudice does harm that can be amplified immeasurably when expressed through behavioral discrimination, hostility, and aggression. A threat to the status or clarity of one's social identity can accentuate conflict and generate group-centrism, zealotry, xenophobia, authoritarian leadership, and persecution of deviance. The most extreme manifestation of harm associated with prejudice is intergroup violence (even genocide), which is facilitated by dehumanization—the perception of out-group members as less than human (often represented as vermin or insects) allows one to treat them as one might treat nonhumans. Clearly, being the victim of extreme intergroup hostility is harmful; it creates disadvantage and stigma and can lead to physical harm and death.

People are, however, resilient and resourceful in protecting themselves against the consequences of stigma—for example, by denying personal experience of stigma or by arguing that knowing one's place and identity is preferable to embarking on a risky strategy of social change. But sometimes enough is enough. Groups are most likely to challenge the status quo when their attainments suddenly drop relative to expectations, when they feel a strong sense of being deprived as a group relative to relevant other groups, and when they believe the status quo is unstable and illegitimate and they can see a feasible strategy to pursue change. The situation is now ripe for protest or revolution—with one key issue being that protest is often dangerous, so supporters of the cause may hold back and hope others will carry the torch and take the risk.

The final section of the chapter explored the role of face-to-face intergroup contact in improving intergroup relations. Contact, even observed and imagined contact, can improve intergroup attitudes, but only if it is pleasant and occurs in the context of shared goals, equal status, and social and institutional support for norms promoting social equality. One challenge from contact is overcoming even well-meaning people's intrinsic anxiety about how contact will play out. A wider challenge for contact is that it needs to be framed in such a way that it is not viewed by either group as eroding group distinctiveness and thus threating social identity, and any favorable outcome of the contact must be generalizable to the out-group as a whole and not viewed as a friendly interpersonal interaction with someone who is not a typical out-group member.

REFERENCES

Aberson, C. L., & Gaffney, A. M. (2009). An integrated threat model of explicit and implicit attitudes. *European Journal of Social Psychology*, *39*, 808–830.

Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology*, *18*, 317–334.

Abrams, D., & Hogg, M. A. (1990). Social identification, self-categorization and social influence. *European Review of Social Psychology*, *1*, 195–228.

Abrams, D., & Hogg, M. A. (2010). Social identity and self-categorization. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE Handbook of Prejudice, Stereotyping and Discrimination* (pp. 179–193). London, UK: Sage.

Abrams, D., Marques, J. M., & Hogg, M. A. (Eds.) (2005). *The Social Psychology of Inclusion and Exclusion*. New York, NY: Psychology Press.

Abrams, D., Marques, J. M., Bown, N., & Henson, M. (2000). Pro-norm and anti-norm deviance within and between groups. *Journal of Personality and Social Psychology*, *78*, 906–912.

Abrams, D., Randsley de Moura, G., Marques, J. M., & Hutchison, P. (2008). Innovation credit: when can leaders oppose their group's norms? *Journal of Personality and Social Psychology*, *95*, 662–678.

Abrams, D., Wetherell, M. S., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: self-categorization and the nature of norm formation, conformity, and group polarization. *British Journal of Social Psychology*, *29*, 97–119.

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. M. (1950). *The Authoritarian Personality*. New York, NY: Harper.

Allport, G. W. (1954). *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Altemeyer, B. (1998). The other "authoritarian personality." *Advances in Experimental Social Psychology*, *30*, 47–92.

Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological Bulletin*, *71*, 319–342.

Bargh, J. A. (1994). The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of Social Cognition* (Vol. 1, 2nd ed., pp. 1–40). Hillsdale, NJ: Erlbaum.

Baron, R. S. (1986). Distraction–conflict theory: progress and problems. *Advances in Experimental Social Psychology*, *20*, 1–40.

Barreto, N. B., & Hogg, M. A. (2017). Evaluation of and support for group prototypical leaders: a meta-analysis of twenty years of empirical research. *Social Influence*, *12*, 41–55.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Blake, R. R., & Mouton, J. S. (1961). Reactions to intergroup competition under win/lose conditions. *Management Science*, *7*, 420–435.

Brewer, M. B. (1991). The social self: on being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Brewer, M. B., & Campbell, D. T. (1976). *Ethnocentrism and Intergroup Attitudes: East African Evidence*. New York, NY: Halsted Press.

Brewer, M. B., & Chen, Y.-R. (2007). Where (who) are collectives in collectivism? Toward conceptual clarification of individualism and collectivism. *Psychological Review*, *114*, 133–151.

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self representation. *Journal of Personality and Social Psychology*, *71*, 83–93.

Burger, J. M. (2009). Replicating Milgram: would people still obey today? *American Psychologist*, *64*, 1–11.

Burnstein, E., & Vinokur, A. (1977). Persuasive argumentation and social comparison as determinants of attitude polarization. *Journal of Experimental Social Psychology*, *13*, 315–332.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, *3*, 14–25.

Cartwright, S., & Schoenberg, R. (2006). Thirty years of mergers and acquisitions research: recent advances and future opportunities. *British Journal of Management*, *17*, S1–S5.

Chen, S., Boucher, H. C., & Tapias, M. P. (2006). The relational self revealed: integrative conceptualization and implications for interpersonal life. *Psychological Bulletin*, *132*, 151–179.

Crandall, C. S. (1994). Prejudice against fat people: ideology and self-interest. *Journal of Personality and Social Psychology*, *66*, 882–894.

Crano, W. D., & Seyranian, V. (2009). How minorities prevail: the context/comparison-leniency contract model. *Journal of Social Issues*, *65*, 335–363.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: the self-protective properties of stigma. *Psychological Review*, *96*, 608–630.

Davies, J. C. (1969). The J-curve of rising and declining satisfaction as a cause of some great revolutions and a contained rebellion. In H. D. Graham & T. R. Gurr (Eds.), *The History of Violence in America: Historical and Comparative Perspectives* (pp. 690–730). New York, NY: Praeger.

Davis, J. H. (1973). Group decision and social interaction: a theory of social decision schemes. *Psychological Review*, *80*, 97–125.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, *51*, 629–636.

Dijksterhuis, A. (2010). Automaticity and the unconscious. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 1, pp. 228–67). Hoboken, NJ: Wiley.

Doosje, B., Branscombe, N., Spears, R., & Manstead, A. (2006). Antecedents and consequences of group-based guilt: the effects of ingroup identification. *Group Processes and Intergroup Relations*, *9*, 325–338.

Dovidio, J. F., & Gaertner, S. L. (1998). On the nature of contemporary prejudice: the causes, consequences, and challenges of aversive racism. In J. L. Eberhardt & S. T. Fiske (Eds.), *Confronting Prejudice: The Problem and the Response* (pp. 3–32). Thousand Oaks, CA: Sage.

Dovidio, J. F., & Gaertner, S. L. (2010). Intergroup bias. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 2, pp. 1084–1121). Hoboken, NJ: Wiley.

Dovidio, J., Glick, P., Hewstone, M., & Esses, V. (Eds.) (2010). *Handbook of Prejudice, Stereotyping and Discrimination*. London, UK: SAGE.

Duckitt, J. (2000). Culture, personality and prejudice. In S. A. Renshon & J. Duckitt (Eds.), *Political Psychology* (pp. 89–107). London, UK: Macmillan.

Duckitt, J., Wagner, C., du Plessis, I., & Birum, I. (2002). The psychological bases of ideology and prejudice: testing a dual process model. *Journal of Personality and Social Psychology*, *83*, 75–93.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*, 573–598.

Ellemers, N. (1993). The influence of socio-structural variables on identity management strategies. *European Review of Social Psychology*, *4*, 27–57.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Festinger, L., Schachter, S., & Back, K. (1950). *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. New York, NY: Harper.

Fielding, K. S., & Hogg, M. A. (2000). Working hard to achieve self-defining group goals: a social identity analysis. *Zeitschrift für Sozialpsychologie*, *31*, 191–203.

Fiedler, F. E. (1964). A contingency model of leadership effectiveness. *Advances in Experimental Social Psychology*, *1*, 149–190.

Fisher, R. J. (1990). *The Social Psychology of Intergroup and International Conflict Resolution*. New York, NY: Springer.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1–74.

Forsyth, D. R. (2006). *Group Dynamics* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Furnham, A. (2003). Belief in a just world: research progress over the past decade. *Personality and Individual Differences*, *34*, 795–817.

Gaertner S. L., & Dovidio, J. F. (2000). *Reducing Intergroup Bias: The Common Ingroup Identity Model*. New York, NY: Psychology Press.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*, 708–724.

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*, 491–512.

Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: development of leader–member exchange (LMX) theory of leadership over 25 years: applying a multi-level multi-domain approach. *Leadership Quarterly*, *6*, 219–247.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology*, *29*, 61–139.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Gurr, T. R. (1970). *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Hamilton, D. L., & Sherman, J. W. (1989). Illusory correlations: implications for stereotype theory and research. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotyping and Prejudice: Changing Conceptions* (pp. 59–82). New York, NY: Springer.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, *103*, 336–355.

Haslam, N. (2006). Dehumanization: an integrative review. *Personality and Social Psychology Review*, *10*, 252–264.

Haslam, N., Loughnan, S., & Kashima, Y. (2008). Attributing and denying humanness to others. *European Review of Social Psychology*, *19*, 55–85.

Haslam, N., Rothschild, L., & Ernst, D. (1998). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*, 113–127.

Hilton, D. J. (2007). Causal explanation: from social perception to knowledge-based causal attribution. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social Psychology: Handbook of Basic Principles* (2nd ed., pp. 232–253). New York, NY: Guilford Press.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations* (2nd ed.). Thousand Oaks, CA: Sage.

Hogg, M. A. (1993). Group cohesiveness: a critical review and some new directions. *European Review of Social Psychology*, *4*, 85–111.

Hogg, M. A. (2005). Social identity and misuse of power: the dark side of leadership. *Brooklyn Law Review*, *70*, 1239–1257.

Hogg, M. A. (2006). Social identity theory. In P. J. Burke (Ed.), *Contemporary Social Psychological Theories* (pp. 111–136). Palo Alto, CA: Stanford University Press.

Hogg, M. A. (2007). Uncertainty-identity theory. *Advances in Experimental Social Psychology*, *39*, 69–126.

Hogg, M. A. (2010). Influence and leadership. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 2, pp. 1166–1207). Hoboken, NJ: Wiley.

Hogg, M. A. (2012a). Social identity and the psychology of groups. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of Self and Identity* (2nd ed., pp. 502–519). New York, NY: Guilford Press.

Hogg, M. A. (2012b). Uncertainty-identity theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (Vol. 2, pp. 62–80). Thousand Oaks, CA: Sage.

Hogg, M. A. (2013). Intergroup relations. In J. DeLamater & A. Ward (Eds.), *Handbook of Social Psychology* (2nd ed., pp. 533–561). New York, NY: Springer.

Hogg, M. A. (2014). From uncertainty to extremism: social categorization and identity processes. *Current Directions in Psychological Science*, *23*, 338–342.

Hogg, M. A. (2015a). Constructive leadership across groups: how leaders can combat prejudice and conflict between subgroups. *Advances in Group Processes*, *32*, 177–207.

Hogg, M. A. (2015b). Social instability and identity-uncertainty: fertile ground for political extremism. In J. P. Forgas, K. Fiedler, & W. D. Crano (Eds.), *Social Psychology and Politics* (pp. 307–320). New York, NY: Psychology Press.

Hogg, M. A., & Abrams, D. (1988). *Social Identifications: A Social Psychology of Intergroup Relations and Group Processes*. London, UK: Routledge.

Hogg, M. A., Adelman, J. R., & Blagg, R. D. (2010). Religion in the face of uncertainty: an uncertainty–identity theory account of religiousness. *Personality and Social Psychology Review*, *14*, 72–83.

Hogg, M. A., & Gaffney, A. M. (2014). Constructing one's social identity: prototype-based comparisons within groups. In Z. Krizan & F. X. Gibbons (Eds.), *Communal Functions of Social Comparison* (pp. 145–174). New York, NY: Cambridge University Press.

Hogg, M. A., & Van Knippenberg, D. (2003). Social identity and leadership processes in groups. *Advances in Experimental Social Psychology*, *35*, 1–52.

Hogg, M. A., Van Knippenberg, D., & Rast, D. E. III. (2012a). Intergroup leadership in organizations: leading across group and intergroup boundaries. *Academy of Management Review*, *37*, 232–255.

Hogg, M. A., Van Knippenberg, D., & Rast, D. E. III. (2012b). The social identity theory of leadership: theoretical origins, research findings, and conceptual developments. *European Review of Social Psychology*, *23*, 258–304.

Hogg, M. A., & Wagoner, J. A. (2017). Normative exclusion and attraction to extreme groups: Resolving identity-uncertainty. In K. D. Williams & S. A. Nida (Eds.), *Ostracism, Exclusion, and Rejection* (pp. 207–223). New York, NY: Routledge.

Hornsey, M. J., & Hogg, M. A. (2000). Assimilation and diversity: an integrative model of subgroup relations. *Personality and Social Psychology Review*, *4*, 143–156.

Inzlicht, M., & Schmader, T. (Eds.) (2011). *Stereotype Threat: Theory, Process, and Application*. New York, NY: Oxford University Press.

Isenberg, D. J. (1986). Group polarization: a critical review. *Journal of Personality and Social Psychology*, *50*, 1141–1151.

Islam, M., & Hewstone, M. (1993). Intergroup attributions and affective consequences in majority and minority groups. *Journal of Personality and Social Psychology*, *65*, 936–950.

Janis, I. L. (1982). *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*, 2nd ed. Boston, MA: Houghton Mifflin.

Jones, E. E., Wood, G. C., & Quattrone, G. A. (1981). Perceived variability of personal characteristics in ingroups and outgroups: the role of knowledge and evaluation. *Personality and Social Psychology Bulletin*, *7*, 523–528.

Jost, J. T., & Hunyadi, O. (2002). The psychology of system justification and the palliative function of ideology. *European Review of Social Psychology*, *13*, 111–153.

Jost, J. T., & Van der Toorn, J. (2012). System justification theory. In P. A. M. van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (Vol. 2, pp. 313–343). London, UK: Sage.

Judd, C. M., & Park, B. (1988). Out-group homogeneity: judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology*, *54*, 778–788.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*, 765–780.

Karau, S. J., & Williams, K. D. (1993). Social loafing: a meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, *65*, 681–706.

Klandermans, B. (2002). How group identification helps to overcome the dilemma of collective action. *American Behavioral Scientist*, *45*, 887–900.

Kruglanski, A. W., Pierro, A., Mannetti, L., & De Grada, E. (2006). Groups as epistemic providers: need for closure and the unfolding of group-centrism. *Psychological Review*, *113*, 84–100.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*, 263–283.

Kurzban, R., & Leary, M. R. (2001). Evolutionary origins of stigmatization: the functions of social exclusion. *Psychological Bulletin*, *127*, 187–208.

Latané, B., Williams, K. D., & Harkins, S. G. (1979). Many hands make light the work: the causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*, 822–832.

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: the sociometer hypothesis. *Journal of Personality and Social Psychology*, *68*, 518–530.

LeBon, G. (1908). *The Crowd: A Study of the Popular Mind*. London, UK: Unwin. (Originally published 1896.)

Leonardelli, G. J., Pickett, C. L., & Brewer, M. B. (2010). Optimal distinctiveness theory: a framework for social identity, social cognition and intergroup relations. *Advances in Experimental Social Psychology*, *43*, 65–115.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*, 275–287.

Levine, J. M. (Ed.) (2013). *Group Processes*. New York, NY: Psychology Press.

Levine, J. M., & Hogg, M. A. (Eds.) (2010). *Encyclopedia of Group Processes and Intergroup Relations*. Thousand Oaks, CA: Sage.

Levine, J. M., & Moreland, R. L. (1994). Group socialization: theory and research. *European Review of Social Psychology*, *5*, 305–336.

Levine, J. M., & Moreland, R. L. (2002). Group reactions to loyalty and disloyalty. *Advances in Group Processes*, *19*, 203–228.

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, *78*, 223–246.

Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annual Review of Psychology*, *58*, 259–289.

Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of the characteristics of in-group and out-group members: empirical evidence and a computer simulation. *Journal of Personality and Social Psychology*, *57*, 165–188.

Lord, R., & Hall, R. (2003). Identity, leadership categorization, and leadership schema. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and Power: Identity Processes in Groups and Organizations* (pp. 48–64). London, UK: Sage.

Mackie, D. M., Maitner, A. T., & Smith, E. R. (2009). Intergroup emotions theory. In T. D. Nelson (Ed.), *Handbook of Prejudice, Stereotyping and Discrimination* (pp. 285–308). New York, NY: Psychology Press.

Macrae, C. N., & Quadflieg, S. (2010). Perceiving people. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 1, pp. 428–463). Hoboken, NJ: Wiley.

Macrae, C. N., Stangor, C., & Hewstone, M. (Eds.). (1996). *Stereotypes and Stereotyping*. New York, NY: Guilford Press.

Major, B. (1994). From social inequality to personal entitlement: the role of social comparisons, legitimacy appraisals and group memberships. *Advances in Experimental Social Psychology*, *26*, 293–355.

Manstead, A. S. R., & Semin, G. R. (1980). Social facilitation effects: mere enhancement of dominant responses? *British Journal of Social and Clinical Psychology*, *19*, 119–136.

Markus, H., & Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Marques, J. M., Abrams, D., & Serôdio, R. (2001). Being better by being right: subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology*, *81*, 436–447.

Marques, J. M., & Páez, D. (1994). The "black sheep effect": social categorisation, rejection of ingroup deviates and perception of group variability. *European Review of Social Psychology*, *5*, 37–68.

Martin, R., & Hewstone, M. (2007). Social influence processes of control and change: conformity, obedience to authority, and innovation. In M. A. Hogg & J. Cooper (Eds.), *The SAGE Handbook of Social Psychology: Concise Student Edition* (pp. 312–332). London, UK: Sage.

Martin, R., & Hewstone, M. (2008). Majority versus minority influence, message processing and attitude change: the source-context-elaboration model. *Advances in Experimental Social Psychology*, *40*, 237–326.

McQueen, A. (2005). A groupthink perspective on the invasion of Iraq. *International Affairs Review*, *14*(2), 53–79.

Miles, E., & Crisp, R. (2014). A meta-analytic test of the imagined contact hypothesis. *Group Processes and Intergroup Relations*, *17*, 3–26.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*, 371–378.

Milgram, S. (1974). *Obedience to Authority*. London, UK: Tavistock.

Miller, C. E. (1985). Group decision making under majority and unanimity decision rules. *Social Psychology Quarterly*, *48*, 51–61.

Moreland, R. L., Hogg, M. A., & Hains, S. C. (1994). Back to the future: social psychological research on groups. *Journal of Experimental Social Psychology*, *30*, 527–555.

Moreland, R. L., & Levine, J. M. (1982). Socialization in small groups: temporal changes in individual-group relations. *Advances in Experimental Social Psychology*, *15*, 137–192.

Moscovici, S. (1976). *Social Influence and Social Change*. London, UK: Academic Press.

Moscovici, S. (1980). Toward a theory of conversion behavior. *Advances in Experimental Social Psychology*, *13*, 202–39.

Mummendey, A., & Otten, S. (1998). Positive–negative asymmetry in social discrimination. *European Review of Social Psychology*, *19*, 107–143.

Mugny, G. (1982). *The Power of Minorities*. London, UK: Academic Press.

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and Social Reality*. Oxford, UK: Blackwell.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the minimal group paradigm: evidence from affective priming tasks. *European Journal of Social Psychology*, *29*, 1049–1071.

Pinto, I. R., Marques, J. M., Levine, J. M., & Abrams, D. (2010). Membership status and subjective group dynamics: who triggers the black sheep effect? *Journal of Personality and Social Psychology*, *99*, 107–119.

Pettigrew, T. F. (1958). Personality and sociocultural factors in intergroup attitudes: a cross-national comparison. *Journal of Conflict Resolution*, *2*, 29–42.

Pettigrew, T. F. (1979). The ultimate attribution error: extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, *5*, 461–476.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783.

Prentice, D. A., Miller, D., & Lightdale, J. R. (1994). Asymmetries in attachment to groups and to their members: distinguishing between common-identity and common-bond groups. *Personality and Social Psychology Bulletin*, *20*, 484–493.

Randsley de Moura, G. R., Leader, T. I., Pelletier, J. P., & Abrams, D. (2008). Prospects for group processes and intergroup relations research: a review of 70 years' progress. *Group Processes and Intergroup Relations*, *11*, 575–596.

Reicher, S. D., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European Review of Social Psychology*, *6*, 161–198.

Reynolds, K. J., Turner, J. C., Haslam, S. A., & Ryan, M. K. (2001). The role of personality and group factors in explaining prejudice. *Journal of Experimental Social Psychology*, *37*, 427–434.

Ridgeway, C. L. (2003). Status characteristics and leadership. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and Power: Identity Processes in Groups and Organizations* (pp. 65–78). London, UK: Sage.

Rokeach, M. (1948). Generalized mental rigidity as a factor in ethnocentrism. *Journal of Abnormal and Social Psychology*, *43*, 259–278.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rubin, M., & Hewstone, M. (1998). Social identity theory's self-esteem hypothesis: a review and some suggestions for clarification. *Personality and Social Psychology Review*, *2*, 40–62.

Runciman, W. G. (1966). *Relative Deprivation and Social Justice*. London, UK: Routledge and Kegan Paul.

Sanders, G. S., & Baron, R. S. (1977). Is social comparison relevant for producing choice shift? *Journal of Experimental Social Psychology*, *13*, 303–314.

Sani, F. (2005). When subgroups secede: extending and refining the social psychological model of schisms in groups. *Personality and Social Psychology Bulletin*, *31*, 1074–1086.

Schachter, S. (1959). *The Psychology of Affiliation: Experimental Studies of the Sources of Gregariousness*. Palo Alto, CA: Stanford University Press.

Schmitt, B. H., Gilovich, T., Goore, N., & Joseph, L. (1986). Mere presence and socio-facilitation: one more time. *Journal of Experimental Social Psychology*, *22*, 242–248.

Schofield, J. W. (1991). School desegregation and intergroup relations: a review of the literature. In G. Grant (Ed.), *Review of Research in Education* (Vol. 17, pp. 335–409). Washington, DC: American Education Research Association.

Sedikides, C., & Strube, M. J. (1997). Self-evaluation: to thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. *Advances in Experimental Social Psychology*, *29*, 209–296.

Sherif, M. (1936). *The Psychology of Social Norms*. New York, NY: Harper.

Sherif, M. (1958). Superordinate goals in the reduction of intergroup conflicts. *American Journal of Sociology*, *63*, 349–356.

Sherif, M. (Ed.) (1962). *Intergroup Relations and Leadership*. New York, NY: Wiley.

Sherif, M. (1966). *In Common Predicament: Social Psychology of Intergroup Conflict and Cooperation*. Boston, MA: Houghton-Mifflin.

Sherif, M., Harvey, O. J., White, B. J., Hood, W., & Sherif, C. (1961). *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. Norman, OK: University of Oklahoma Institute of Intergroup Relations.

Sidanius, J., & Pratto, F. (1999). *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression*. New York, NY: Cambridge University Press.

Simon, B., & Brown, R. J. (1987). Perceived intragroup homogeneity in minority-majority contexts. *Journal of Personality and Social Psychology*, *53*, 703–711.

Smith, E. R., Segar, C. R., & Mackie, D. M (2007). Can emotions be truly group level? Evidence regarding four conceptual criteria. *Journal of Personality and Social Psychology*, *93*, 431–446.

Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). A terror management theory of social behavior: the psychological functions of self-esteem and cultural worldviews. *Advances in Experimental Social Psychology*, *24*, 93–159.

Stangor, C. (2016). *Social Groups in Action and Interaction*, 2nd ed. New York, NY: Routledge.

Stasser, G., & Titus, W. (2003). Hidden profiles: a brief history. *Psychological Inquiry*, *14*, 304–314.

Staub, E. (2010). *The Panorama of Mass Violence: Origins, Prevention, Healing and Reconciliation*. New York, NY: Oxford University Press.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: the psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing Prejudice and Discrimination* (pp. 23–46). Mahwah, NJ: Erlbaum.

Stürmer, S., & Simon, B. (2004). Collective action: towards a dual-pathway model. *European Review of Social Psychology*, *15*, 59–99.

Swann, W. B., Jetten, J., Gomez, A., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: a theory of identity fusion. *Psychological Review*, *119*, 441–456.

Swim, J. T., Aikin, K., Hall, W., & Hunter, B. A. (1995). Sexism and racism: old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*, 199–214.

Tajfel, H. (1959). Quantitative judgement in social perception. *British Journal of Psychology*, *50*, 16–29.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*, 96–102.

Tajfel, H., Billig, M., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, *1*, 149–177.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 7–24). Chicago, IL: Nelson-Hall.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology*, *9*, 507–533.

Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, *27*, 585–600.

Turner, J. C. (1991). *Social Influence*. Milton Keynes, UK: Open University Press.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, UK: Blackwell.

Turner, J. C., Wetherell, M. S., & Hogg, M. A. (1989). Referent informational influence and group polarization. *British Journal of Social Psychology*, *28*, 135–147.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. *Advances in Experimental Social Psychology*, *25*, 115–191.

Van den Bos, K. (2009). Making sense of life: the existential self trying to deal with personal uncertainty. *Psychological Inquiry*, *20*, 197–217.

Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, *134*, 504–535.

Verkuyten, M. (2006). Multiculturalism and social psychology. *European Review of Social Psychology*, *17*, 148–184.

Wagoner, J. A., & Hogg, M. A. (2016). Normative dissensus, identity-uncertainty, and subgroup autonomy. *Group Dynamics: Theory, Research, and Practice*, *20*, 310–322.

Walker, I., & Pettigrew, T. F. (1984). Relative deprivation theory: an overview and conceptual critique. *British Journal of Social Psychology*, *23*, 301–310.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961–977.

Wilder, D. A. (1984). Intergroup contact: the typical member and the exception to the rule. *Journal of Experimental Social Psychology*, *20*, 177–194.

Williams, K. D. (2001). *Ostracism: the Power of Silence*. New York, NY: Guilford Press.

Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: the effects of expectations of co-worker performance. *Journal of Personality and Social Psychology*, *61*, 570–581.

Williams, K. D., Karau, S. J., & Bourgeois, M. (1993). Working on collective tasks: social loafing and social compensation. In M. A. Hogg & D. Abrams (Eds.), *Group Motivation: Social Psychological Perspectives* (pp. 130–148). London, UK: Harvester Wheatsheaf.

Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: a meta-analytic review of social influence processes. *Psychological Bulletin*, *115*, 323–345.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*, 73–90.

Yap, S., Ji, L., & Hong, E. (2017). Culture and cognition. In J. T. Wixted & S. Thompson-Schill (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 3: *Language and Thought*). Hoboken, NJ: Wiley.

Yukl, G. (2010). *Leadership in Organizations*, 7th ed. Upper Saddle River, NJ: Pearson Education.

Yzerbyt, V., & Demoulin, S. (2010). Intergroup relations. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 2, pp. 1024–1083). Hoboken, NJ: Wiley.

Yzerbyt, V., Judd, C. M., & Corneille, O. (Eds.) (2004). *The Psychology of Group Perception: Perceived Variability, Entitativity, and Essentialism*. New York, NY: Psychology Press.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 269–274.

Zhang, S., & Hunt, J. S. (2008). The stereotype rebound effect: universal or culturally bounded process? *Journal of Experimental Social Psychology*, *44*, 489–500.

Zimbardo, P. G. (1969). The human choice: individuation, reason, and order vs. deindividuation, impulse, and chaos. In W. J. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation* (pp. 237–307). Lincoln, NE: University of Nebraska Press.

CHAPTER 15

Explicit and Implicit Emotion Regulation

LAURA MARTIN BRAUNSTEIN AND KEVIN N. OCHSNER

INTRODUCTION

Everyday we use emotion regulation strategies to soften the blow of negative experiences or to reduce the allure of temptations. Taking control over our emotions is important for health and for achieving long-term goals. Although a variety of emotion regulation strategies exist, research has focused on explicit forms of emotion regulation—strategies that we deliberately pursue and that require cognitive control and effort—and most of these studies have been conducted in healthy adults. However, owing to their reliance on conscious initiation and cognitive control processes, explicit emotion regulation strategies are not well suited for some populations and situations. For that reason, it is important to understand other forms of emotion regulation that do not depend on intact and fully developed cognitive control abilities and to determine which forms of emotion regulation may be most beneficial for a given population. The goal of this chapter is to describe and categorize the full range of emotion regulation strategies, their neural systems, and their efficacy for different clinical populations and across the life span.

Toward these ends, this chapter offers a multilevel framework that can account for a wide range of emotion regulation strategies, including explicit and implicit forms of regulation. In building the framework, we drew on behavioral and brain data to understand the affective and cognitive processes involved in emotion regulation. As we illustrate, this framework uses three levels of analysis—behavior, psychological processes, and neural systems—to describe and differentiate four basic classes of emotion regulation strategies. The critical aspect of the framework is its description of the psychological processes underlying emotion regulation using two fundamental dimensions: (1) the nature of the emotion regulation goal (ranging from implicit to explicit) and (2) the nature of the emotion change process (ranging from more automatic to more controlled). Describing the core process dimensions of emotion regulation using this scheme allows us to make strong inferences about the neural systems that subserve each dimension, creating a comprehensive neurofunctional framework for emotion regulation.

To elaborate this framework, the remainder of the chapter is divided into three sections. The first section defines the emotion regulation framework and illustrates how different combinations of emotion regulation goals and emotion change processes yield the four classes of emotion regulation strategies: explicit controlled, implicit automatic, implicit controlled, and explicit automatic.

The second section reviews neural and behavioral features of each of the four classes of emotion regulation. The third section identifies the advantages of the framework and suggests directions for future work, including possible applications of the different classes of strategies to clinical populations and development/aging.

DEFINING AND CLASSIFYING EMOTION REGULATION STRATEGIES

Before delving into the details of the framework, it is useful to first briefly describe and define emotion and emotion regulation. To be as inclusive as possible, we take the perspective that emotions are sets of valenced behavioral, cognitive, physiological, and experiential responses to stimuli that reflect appraisals of a stimulus's relevance to current goals, needs, or values (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Moors, Ellsworth, Scherer, & Frijda, 2013; Panksepp, 2003; Scherer, Schorr, & Johnstone, 2001; Schultz, Dayan, & Montague, 1997). The neural systems involved in generating emotion responses have been described in detail elsewhere (for review, see Kober et al., 2008; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Vytal & Hamann, 2010), but we highlight the key features here.

The generation of emotion is supported by several neural systems, and the specific regions that are recruited depend on the stimulus and one's emotional response. Three regions have been most commonly implicated in studies of emotion and its regulation, including the amygdala, which has been implicated in detecting, encoding, and initiating responses to goal-relevant stimuli (Cunningham & Brosch, 2012; Davis & Whalen, 2001; Janak & Tye, 2015; Phelps, 2006); the ventral striatum, which has been implicated in learning the value of stimuli (Delgado, 2007; O'Doherty, 2004); and the insula, which may support awareness of affective body states (Craig, 2009; Critchley, Wiens, Rotshtein, Öhman, & Dolan, 2004; Zaki, Davis, & Ochsner, 2012).

Emotion regulation occurs when we modify the nature, intensity, or duration of our emotional response in line with emotion regulatory goals. Strategies for regulating emotion can be distinguished in many ways (e.g., see Gross, 1998a, 1998b; Gross & Barrett, 2011; Tice & Bratslavsky, 2000; Webb, Schweiger Gallo, Miles, Gollwitzer, & Sheeran, 2012); here we focus on a high-level classification scheme that differentiates classes of emotion regulation strategies using two orthogonal dimensions (see Figure 15.1): (1) the nature of the goal to regulate and (2) the nature of the emotion change process.

The first dimension (see *y*-axis of Figure 15.1) describes whether the goal to regulate is explicit or implicit. For present purposes, goals are defined as mental representations of potential states of behavior or experience (e.g., goal to feel happier), which can be consciously held (explicit) or nonconscious (implicit) (Bargh & Williams, 2007; Hassin, 2013; Hassin, Aarts, Eitam, Custers, & Kleiman, 2009; Hassin, Ochsner, & Trope, 2010). In this view, explicit goals involve a conscious desire to change one's emotions and/or the conscious expectation that change will occur whereas implicit goals do not involve the conscious desire to change emotional responding. Although it may seem counterintuitive to think of goals as implicit, our conceptualization is guided by the literature on automatic goal pursuit, which demonstrates that goals can be activated and shape behavior outside of conscious awareness (Custers & Aarts, 2010). Some implicit goals are chronically and continuously active,

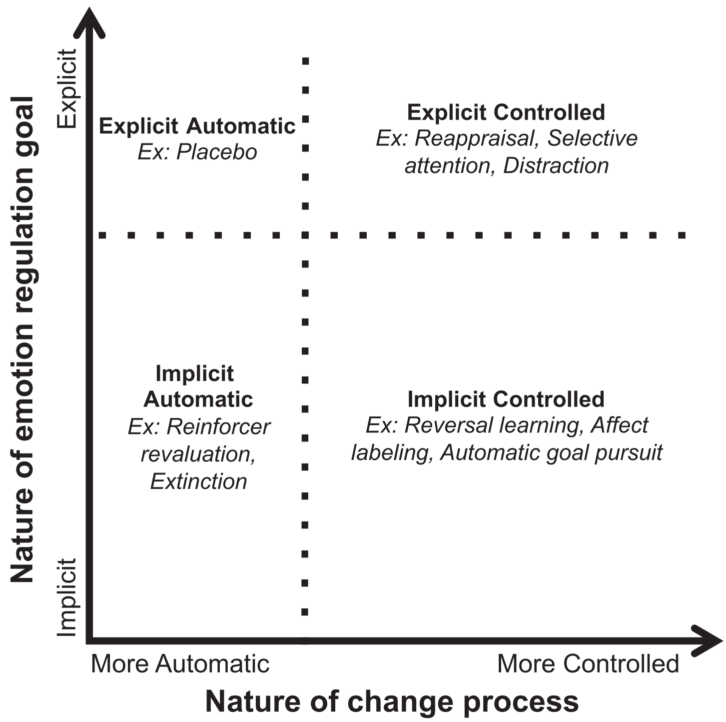


Figure 15.1 A classification scheme for emotion regulation strategies based on two orthogonal dimensions: (1) the nature of the goal to regulate (implicit vs. explicit) and (2) the nature of the emotion change process (automatic vs. controlled).

such as the basic goal to maintain accurate representations of the values of stimuli in the environment. Others are transient and context specific, such as when a goal is primed nonconsciously by cues present in a specific situation (Shidlovski & Hassin, 2011) or when the performance of a task—whose overt goal is not emotion-related—has the incidental effect of altering emotional responses (Lieberman et al., 2007).

The second dimension (see *x*-axis of Figure 15.1) describes the nature of the emotion change process—the means by which an emotional response is altered. At one end of this dimension are (primarily if not exclusively) automatic change processes that operate without conscious monitoring or awareness that change is occurring. Automatic change processes may require attention or effort, but the attention or effort expended is not perceived as related to the subsequent change in affect (Hassin, 2013). At the other end of this dimension are (primarily if not exclusively) controlled change processes that typically involve the effortful and/or attention-demanding representation of task goals in working memory and conscious monitoring of regulatory progress, and may involve other cognitive control processes including response inhibition or selection (Badre & Wagner, 2004; Miller & Cohen, 2001; Ochsner, Silvers, & Buhle, 2012).

Although this two-dimensional framework highlights that the nature of both the goals and the change processes can vary by degrees along a continuum, qualitatively different forms of regulation arise from different combinations of processes and goals.

Critically, explicit and implicit emotion regulation are most cleanly dissociated as occupants of two distinct quadrants along the major diagonal within Figure 15.1, where explicit emotion regulation involves intentional goals initiating regulation via controlled change processes, and implicit emotion regulation involves implicit goals initiating regulation via nonconscious change processes. Beyond this core distinction, the framework also offers a means of understanding two other strategies that lie in the remaining two quadrants along the minor diagonal of the figure. As discussed next, we call them hybrid strategies because they share features of explicit strategies on one dimension and features of implicit strategies on the other.

FOUR CLASSES OF EMOTION REGULATION

With these definitions in place, we can now review extant behavioral and brain data to characterize the four classes of emotion regulation. For each class, we first describe prototypical exemplar strategies (i.e., the most commonly studied strategies), their core underlying neural systems generally (see Figure 15.2), and finally their application in clinical populations and across the life span.

The latter is particularly important because mapping the behavioral characteristics and neural mechanisms for different kinds of emotion regulation has great translational value. For example, we can apply the framework to identify the neural origins of normal and abnormal variability in emotion regulation ability. In the sections that follow, we describe the behavioral and neural mechanisms of each class of emotion regulation and consider how effective each class is for different clinical populations or age groups.

Explicit-Controlled Emotion Regulation

Explicit-controlled emotion regulation strategies are perhaps the best understood because they are the most commonly studied of all regulatory strategies. Such strategies are engaged deliberately by individuals who are aware that they are regulating and depend

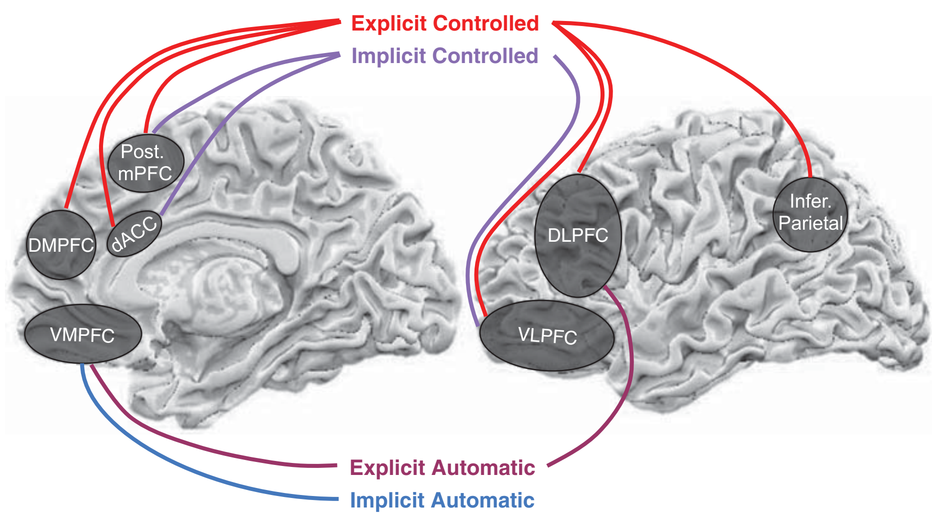


Figure 15.2 Brain regions hypothesized to underlie the emotion regulation processes as illustrated in Figure 15.1. Color version of this figure is available at http://onlinelibrary.wiley.com/book/10.1002/9781119170174.

on prefrontal systems that support various kinds of cognitive control processes. The most commonly studied explicit regulation strategies are selective attention/distraction and reappraisal, which involve changing how one attends to or appraises a stimulus, respectively (Ochsner & Gross, 2005; see Chapter 9 in this volume).

Prototypical Exemplars

Selective Attention and Distraction. Anyone who has ever closed their eyes during a gruesome scene in a horror movie can attest that what we attend to affects how we feel. When we effortfully control the focus of our attention to change how we feel, we are using attentional control strategies to explicitly regulate our emotion. In general, two kinds of such strategies have been studied: selective attention and distraction (Ochsner & Gross, 2005). Selective attention involves focusing on certain features of an affective stimulus and ignoring other features, whereas distraction involves shifting the focus of attention entirely away from the stimulus. Selective attention studies take two main forms—those that ask participants to focus on and evaluate emotional aspects of the stimuli (e.g., judge the expression of a face) and those that ask them to attend to and evaluate nonemotional features (e.g., indicate the gender of a face) (Ochsner & Gross, 2005). In typical distraction paradigms, participants are presented with emotional stimuli, and they are asked to simultaneously engage in an attention-demanding secondary task (the distractor), such as keeping in mind a string of letters over a delay (McRae et al., 2010) or performing mental arithmetic (Kanske, Heissler, Schönfelder, Bongers, & Wessa, 2010; Van Dillen, Heslenfeld, & Koole, 2009).

There is evidence that the brain regions involved in regulation via selective attention and distraction differ, but understanding these differences is complicated by the fact that the two strategies are studied using different tasks, have been used to regulate different kinds of stimuli, and have assessed emotion changes using different dependent measures (e.g., brain activation versus self-reports of emotion experience).

On one hand, selective attention studies typically use low-arousal images of emotional facial expressions, use amygdala activation as the marker of emotion change, and rarely assess experience or other behavioral indicators of emotion. In such studies, attending to a stimulus's emotional features might reduce amygdala activity (Critchley et al., 2000; Taylor, Phan, Decker, & Liberzon, 2003)—a finding taken as evidence that attention to emotion can diminish emotional responding—whereas attending to nonemotional aspects does not impact amygdala response (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Vuilleumier, Armony, Driver, & Dolan, 2001; Winston, O'Doherty, & Dolan, 2003)—a finding taken as evidence that emotional responding is automatic and unaffected by attention. Although the reasons for these discrepant findings are not yet clear, factors such as the degree of cognitive load or semantic processing required by a given task may be key (Ochsner & Gross, 2005; Pessoa, 2009). Although it is counterintuitive, a meta-analysis found that the amount of attentional processing affects amygdala activity, with passive viewing of affective stimuli leading to greater activation than directed/instructed viewing (Costafreda, Brammer, David, & Fu, 2008).

On the other hand, distraction studies largely use high-arousal aversive stimuli, especially those associated with physical pain; sometimes measure behavior or experience; and, as evidence that regulation occurred, examine both the engagement of brain regions involved in cognitive control and the modulation of regions involved in pain and affect. Typically, such studies report activation of prefrontal and parietal regions including the dorsolateral prefrontal cortex (DLPFC) (Kanske et al., 2010; Van Dillen et al., 2009), dorsomedial PFC (DMPFC) (Kanske et al., 2010; McRae et al., 2010), dorsal anterior cingulate cortex (dACC) (Kanske et al., 2010; McRae et al., 2010), and parietal cortex (Kanske et al., 2010; McRae et al., 2010) along with modulation of the amygdala, anterior insula, and/or mid-cingulate regions that support nociception (Wager et al., 2013). Although these studies highlight the potential role of prefrontal regions in the regulation of affect-generating systems, clear interpretations are complicated by the fact that activations could reflect performance of the secondary distraction task, conflict between responses generated by the stimulus and the secondary task, or some combination of both.

Reappraisal. When we change our emotional response to a stimulus by cognitively altering the way we appraise its meaning, we are using reappraisal. Reappraisal can be used to increase and decrease the intensity, or change the quality, of various kinds of emotional responses in accordance with regulation goals (Gross, 1998a; Ochsner et al., 2004). To date, most work has examined the use of reappraisal to decrease responses to aversive visual stimuli (e.g., photographic images). A canonical version of such a task presents aversive pictures and asks participants to reappraise them by reinterpreting depicted scenes in a less negative way. For example, when seeing a picture of an injured man in a hospital, one could describe him as being of a strong constitution and healing quickly.

Reappraisal is thought to rely on domain-general cognitive control processes and neural systems to be described. Consistent with that theory, working memory capacity may predict behavioral indicators of reappraisal success (Schmeichel, Volokhov, & Demaree, 2008), and better cognitive performance has been associated with greater decreases in amygdala activity during reappraisal (Winecoff et al., 2011). Several reviews (Green & Malhi, 2006; Ochsner & Gross, 2005, 2008; Ochsner et al., 2012) and two meta-analyses (Buhle et al., 2014; Diekhof, Geier, Falkai, & Gruber, 2011) consistently show that the act of reappraising activates three kinds of systems implicated in cognitive control. These include: DLPFC and parietal cortex, thought to support working memory for, and selective attention to, reappraisal-relevant information; ventrolateral PFC (VLPFC), thought to support selection of stimulus-appropriate reappraisals and inhibition of prepotent evaluations; and posterior portions of DMPFC and dACC, thought to be important for monitoring the success of reappraisal. When the reappraisal goal is to diminish affective responding, activation may be stronger in right lateral prefrontal regions commonly involved in the inhibition of prepotent response. When the goal is to enhance affective responses, activation may be stronger in anterior portions of the DMPFC that may support elaboration of and attributions about the affective meaning of stimuli. Although no studies have compared reappraisal of different kinds of stimuli or emotions directly, there is evidence that the regions modulated by reappraisal vary as a function of both variables (Ochsner et al., 2012; Woo, Roy, Buhle & Wager, 2015). In general, heightened activity in goal-relevant PFC regions may be correlated with diminished activity in brain regions involved in emotional responding and computing affective values more generally, including the amygdala for threat-related stimuli and the ventral striatum for stimuli with appetitive value

(Buhle et al., 2014; Ochsner et al., 2012). There seems to be an inverse relationship between PFC activity related to cognitive control and activity in brain regions involved in emotion, and this relationship is the underlying neural mechanism of reappraisal.

Core Neural Systems

What we know about the neural mechanisms that support explicit-controlled emotion regulation comes largely from human imaging studies, because there is no animal work and scant lesion work. Animal models of explicit-controlled regulation are somewhat impractical, because this form of regulation depends on explicit goals and complex cognitive processes. We know of two lesion studies in humans: a case study of a patient with a left frontal stroke showing she was unable to generate reappraisals spontaneously (Salas, Gross, Rafal, Viñas-Guasch, & Turnbull, 2013) and a study of patients with focal frontal unilateral lesions who were slower to generate reappraisals (Salas, Gross, & Turnbull, 2014).

Explicit-controlled emotion regulation strategies are supported by increased engagement of prefrontal and cingulate cortical brain systems involved in cognitive control, with different combinations of systems recruited depending on the specific strategy (Buhle et al., 2014; Green & Malhi, 2006; Ochsner & Gross, 2005; Ochsner et al., 2012). Importantly, the recruitment of these prefrontal systems fosters changes in activity in brain regions involved in emotional responding, including the amygdala, striatum, and insula (Buhle et al., 2014; Green & Malhi, 2006; Ochsner & Gross, 2007; Ochsner et al., 2012).

Our understanding of the neural systems for explicit emotion regulation builds on prior cognitive neuroscience research documenting the importance of the PFC and ACC in effortful forms of goal-directed cognition and action, in general. In this work, cognitive control is thought to be implemented through the influence of domain-general PFC-based control systems on posterior and subcortical systems that represent specific kinds of sensory or mnemonic information (Botvinick, Cohen, & Carter, 2004; Knight, Staines, Swick, & Chao, 1999; Miller & Cohen, 2001; Smith & Jonides, 1999). This influence is possible due to the intrinsic and extrinsic anatomical connections of the PFC/parietal control regions (Miller & Cohen, 2001; Miller & D'Esposito, 2005) that allow them to influence processing in other parts of the brain, which is necessary for goal-directed control of behavior. For example, specific anatomical and functional frontal-subcortical networks have been identified that contribute to control of motor responses, such as the response inhibition that is required during the stop-signal task (Aron, Behrens, Smith, Frank, & Poldrack, 2007). Imaging research on emotion regulation has shown that a similar dynamic is at play during explicit emotion regulation, with control systems modulating processing in brain regions that generate affective responses (Green & Malhi, 2006; Ochsner & Gross, 2005, 2008; Ochsner et al., 2012). As described next, three systems have been most strongly implicated in the explicit regulation of affective responses: DLPFC and parietal cortex; VLPFC; and dACC and adjacent DMPFC.

DLPFC and Parietal Cortex. In general, the DLPFC has been implicated in the active maintenance of task rules/goals as well as sensory information relevant to these goals, including stimuli and their contingent relationships with behavioral responses (Badre & Wagner, 2004; Curtis & D'Esposito, 2003; Miller & Cohen, 2001). These control processes contribute to emotion regulation, and the DLPFC is involved in distraction (Kanske et al., 2010; Van Dillen et al., 2009) and reappraisal (Buhle et al., 2014; Diekhof et al., 2011; Green & Malhi, 2006; Ochsner & Gross, 2005, 2008; Ochsner et al., 2012). The DLPFC likely supports the process of maintaining the goal of distracting oneself from or generating new interpretations for affective stimuli. Distraction and reappraisal both involve holding and manipulating information in working memory and controlling the focus of attention, which are thought to depend on a DLPFC–parietal control network (Miller & Cohen, 2001; Wager & Smith, 2003; Wager, Jonides, & Reading, 2004).

VLPFC. The VLPFC is involved in the selection of desired responses and inhibition of inappropriate ones (Aron, Robbins, & Poldrack, 2004; Cohen & Lieberman, 2010). Reappraisal requires selecting desired evaluations of stimuli and inhibiting prepotent reactions. The VLPFC is recruited during reappraisal (Buhle et al., 2014; Diekhof et al., 2011; Green & Malhi, 2006; Ochsner & Gross, 2005, 2008; Ochsner et al., 2012). Consistent with the functional imaging data, gray matter volume in the right VLPFC predicts both reappraisal ability and performance on the stop-signal task, a common measure of response inhibition (Tabibnia et al., 2011).

DACC and DMPFC. The dACC and adjacent DMPFC monitor conflicts between the intended and actual behavioral outcomes and signal when appropriate adjustments in control are needed (Badre & Wagner, 2004; Botvinick, Braver, Barch, Carter, & Cohen, 2001, 2004; Carter & Van Veen, 2007). These functions are important in explicit-controlled regulation. The dACC is involved in distraction (Kanske et al., 2010; McRae et al., 2010) and has been related to successful reappraisal in some studies (Ochsner, Bunge, Gross, & Gabrieli, 2002). The DMPFC is engaged during reappraisal and likely supports monitoring and reflection on one's emotional states (Buhle et al., 2014; Ochsner et al., 2012).

Affective Systems. Explicit strategies modulate activity in affective brain systems, chiefly the amygdala, ventral striatum, and insula. Amygdala activity is decreased during selective attention to (negative) emotional features (Critchley et al., 2000; Taylor et al., 2003) but not nonemotional features (Anderson et al., 2003; Vuilleumier et al., 2001; Winston et al., 2003). In negative emotion regulation, amygdala activity is consistently reduced by distraction (Kanske et al., 2010; McRae et al., 2010; Van Dillen et al., 2009) and reappraisal (Buhle et al., 2014; Diekhof et al., 2011; Green & Malhi, 2006; Ochsner & Gross, 2005, 2008; Ochsner et al., 2012). Insula activity is reduced by distraction from pain (Wager et al., 2013) and sometimes by distraction from negative emotion (Van Dillen et al., 2009). Regulation of positive emotion using explicit strategies is less studied, but ventral striatum activity has been reduced by distraction (Delgado, Gillis, & Phelps, 2008; Martin & Delgado, 2011) and reappraisal (Kober, Kross, Mischel, Hart, & Ochsner, 2010a; Kober et al., 2010b).

Implicit-Automatic Emotion Regulation

Implicit-automatic emotion regulation refers to strategies initiated by implicit goals to change affect that are implemented by processes that operate fairly automatically. The regulatory goals could either be chronic, such as the continuously operating goal to accurately represent and when necessary change or update the value of affective stimuli (LeDoux, 2012); or they can be implicit and transient, such as when as the performance of a given task has the incidental effect of changing the way one attends to affective information (Lieberman et al., 2007). In both cases, emotional responses are altered by processes that operate without conscious monitoring or awareness. Prototypical exemplars of implicit regulation include forms of affective learning, specifically, extinction and reversal learning.

Prototypical Exemplars

Extinction. In extinction, an individual learns via experience that the affective value of a stimulus has changed and updates affective responses to the stimulus accordingly. A classic example is fear extinction (Stjepanovic, 2017). When a stimulus such as a tone has been paired repeatedly with shock, it becomes a conditioned stimulus that evokes affective responses associated with the anticipation of shock, such as fear and freezing. When the tone ceases to be paired with the shock, a new value for the tone is learned and fear responses to the tone are diminished or extinguished. Animal and human research has demonstrated the involvement of the ventromedial PFC (VMPFC) in the retention of these newly learned values (Hartley & Phelps, 2009; Milad & Quirk, 2002; Phelps, Delgado, Nearing, & Ledoux, 2004; Quirk, Russo, Barron, & Lebron, 2000), and individual differences in the thickness of the VMPFC in humans are correlated with retention (Milad et al., 2005).

Reinforcer Reevalution. In reinforcer devaluation and revaluation studies, the positive outcome associated with presentation of a stimulus is either increased or reduced, and hence its affective value changes. Responses to it must be regulated and mental representations of value must be updated. In classic devaluation studies in animals, the value of a food stimulus is reduced either by pairing it with illness or by allowing the animal to eat it until satiety, and this devaluation diminishes responses in the VMPFC (Kerfoot, Agarwal, Lee, & Holland, 2007; Murray, O'Doherty, & Schoenbaum, 2007). Similar results have been found in human imaging studies (Gottfried, O'Doherty, & Dolan, 2003).

Core Neural Systems

Our understanding of the neural systems involved in implicit-automatic emotion regulation is derived from a combination of animal research and human neuroimaging and lesion studies. The supporting brain regions—foremost the VMPFC/orbitofrontal cortex (OFC)—often are thought of as learning systems that track—and, when needed, update—representations of the affective value of stimuli to support contextually appropriate responses to them. For simplicity, we use the term "VMPFC" throughout the chapter.

VMPFC. The core process in implicit-automatic regulation is updating the affective value of a stimulus based on the current context, including goals, wants, and needs. These kinds of contextualized value computations have been studied extensively in both human and animal research (for review, see Delgado, Olsson, & Phelps, 2006; Holland & Gallagher, 2004; Murray et al., 2007; Pickens & Holland, 2004; Quirk & Mueller, 2007), and this work provides insights into their role in implicit regulation. Meta-analyses of imaging work have shown that the VMPFC is the primary, but not only, region whose activity tracks the subjective value of food, money, or various goods (Chib, Rangel, Shimojo, & O'Doherty, 2009; Elliott, Agnew, & Deakin, 2008; Plassmann, O'Doherty, & Rangel, 2007). Human and animal work has demonstrated that VMPFC responses are diminished to stimuli that have been devalued (Gottfried et al., 2003; Kerfoot et al., 2007).

Lesions studies also support the role of the VMPFC in representing subjective value/preferences based on the current context. VMPFC lesions impair both the initial learning (Camille, Tsuchida, & Fellows, 2011b) and the expression of simple preferences for stimuli (Fellows & Farah, 2007), render preference hierarchies unstable (Camille, Griffiths, Vo. Fellows, & Kable, 2011a), and disrupt evaluation of social stimuli such as facial expressions (Heberlein, Padon, Gillihan, Farah, & Fellows, 2008; Hornak, Rolls, & Wade, 1996) and stereotyped groups (Milne & Grafman, 2001). Additionally, individuals with VMPFC lesions exhibit behaviors that suggest they have an impaired ability to use social contexts and norms to guide their behavior. For instance, they show general disruptions of social behavior and emotional experience (Hornak et al., 2003), including teasing strangers in inappropriate ways (Beer, Heerey, Keltner, Scabini, & Knight, 2003) and failing to recognize when others commit a faux pas (Stone, Baron-Cohen, & Knight, 1998). Finally, animal research (for review, see Schoenbaum, Takahashi, Kiu, & McDannald, 2011) dovetails with these findings by demonstrating that the VMPFC is necessary for in-the-moment value estimates that rely on current information about feedback contingencies, as rats with lesions to the VMPFC expressed only stored preferences for stimuli and did not prefer stimuli that were newly associated with rewards, whereas controls made both distinctions (Jones et al., 2012; Schoenbaum et al., 2011).

The connectivity of the VMPFC is well suited for this role, as it is reciprocally interconnected with regions that: (a) identify what stimuli are present—including lateral temporal regions representing perceptual features of stimuli (Öngür & Price, 2000); (b) indicate the initial valuation of these stimuli and their features—including key connections with the amygdala (Amaral & Price, 1984) and ventral striatum (Haber & Knutson, 2010; Öngür & Price, 2000) that provide information about the relevance of stimuli to affective goals; and (c) provide important contextual information that may constrain the range of appropriate responses—including connections with medial temporal regions that may have encoded specific prior experiences with stimuli (Price & Drevets, 2009), subcortical and cortical (e.g., insula) regions representing current motivational (e.g., hunger) and other body states (e.g., pain) (Cavada, Tejedor, Cruz-Rizzolo, & Reinoso-Suárez, 2000; Price, 2007), and the lateral prefrontal and cingulate regions that represent current task goals (Cavada et al., 2000; Price, 2007).

This diverse set of connections allows the VMPFC to compute the current affective value of a stimulus, given one's current motivational state, current goals, and prior history with the stimulus.

Importantly, and as unpacked for select prototypical examples in the next paragraph, the VMPFC works together with the amygdala and striatum to compute an expectancy about the affective value of a stimulus in the current context.

In humans, the VMPFC is thought to compute the current subjective value (often referred to as affective value) of a stimulus by aggregating inputs from many different brain regions (Roy, Shohamy, & Wager, 2012). Subjective value by its very definition must take into account both internal factors, including goals, wants, and needs, and external information, including context, experienced outcomes or feedback, and environmental constraints. In affective neuroscience, subjective value is synonymous with affective value, and these terms often are used interchangeably. Human neuroimaging paradigms designed to measure the neural systems that correlate with subjective value for food, money, or various goods have identified the VMPFC/OFC as the primary brain region whose activity tracks with reported values (Chib et al., 2009; Elliott et al., 2008; Plassmann et al., 2007). Animal research has demonstrated that the OFC conducts online estimates of the value of a stimulus in the moment using relevant contextual information (Jones et al., 2012), in particular when the value is dependent on outcomes associated with the stimulus (Schoenbaum et al., 2011).

Although this section is on implicit-automatic regulation, it is important to point out that activity in the VMPFC can be altered by both implicit *and* explicit goals. Research in humans has shown that explicit goals change the value signals tracked by VMPFC. Take, for example, the goal to diet, which starts as an explicit goal but may become more implicit with experience. In dieters, activity in VMPFC correlates with evaluations of how tasty *and* healthy foods are, whereas in nondieters, VMPFC activity is associated only with taste information (Hare, Camerer, & Rangel, 2009).

Finally, studies of patients with lesions of the VMPFC reveal the importance of this region in linking behaviors and affective responses to situational and social contexts and in promoting contextually appropriate responding. Individuals with OFC lesions exhibit inappropriate behavior—for instance, disclosing intimate personal details to and teasing strangers (Beer et al., 2003) and failing to recognize when others commit a faux pas (Stone et al., 1998). These altered behaviors are present even when the behavior assessed is implicit—patients with VMPFC lesions do not exhibit implicit gender stereotypes typically observed in controls (Milne & Grafman, 2001).

Implicit-Controlled Regulation

Implicit-controlled regulation occurs when there is an implicit goal to regulate emotion and emotion experience is changed by controlled processes. Regions of the lateral PFC and dACC involved in selection and inhibition and in conflict monitoring have been implicated in implicit-controlled regulation. The strategies that fall into the implicit-controlled class of emotion regulation are somewhat diverse and include affect labeling, emotional Stroop and emotional go/no go tasks, reversal learning, and automatic goal pursuit.

Prototypical Exemplars

Implicit-controlled regulation occurs when response selection and/or inhibition processes are engaged and the activation of these control processes dampens affective responding in the absence of an intentional goal to regulate. Often, implicit-controlled regulation occurs as a by-product of performing another task. We have identified several tasks that fit the purview of implicit-controlled regulation: affect labeling and the emotional Stroop and emotional go/no go tasks.

Affect Labeling. Affect labeling, the act of selecting a brief written description for an emotional stimulus, reduces behavioral and neural affective responding and has been described to be an incidental emotion regulation process (Berkman & Lieberman, 2009). Affect labeling is akin to a pared-down version of writing about emotional experiences, which improves mental and physical health (Pennebaker, 1997). A typical affect labeling paradigm consists of two conditions: one in which the participant assigns an affective label to an affective image, and a control condition in which the participant matches to another image rather than label or give a nonaffective label. Initial studies tested the effect of labeling on neural responses to emotional facial expressions and observed that labeling decreased amygdala responding and recruited right VLPFC (Hariri, Bookheimer, & Mazziotta, 2000; Lieberman et al., 2007). Follow-up studies have extended these results by demonstrating that labeling reduces self-reported emotion and by showing that the effect is not limited to faces—affect labeling reduces self-reports of negative and positive emotion elicited by complex emotional scenes (Lieberman, Inagaki, Tabibnia, & Crockett, 2011). Evidence that labeling has implicit regulatory effects comes from a series of studies that asked participants to predict or reflect on the anticipated/experienced effects of labeling on emotional responding. Participants consistently reported that labeling would or did increase emotion—the opposite of the observed effect—suggesting that the regulatory effects of affect labeling are implicit and not due to a conscious goal (Lieberman et al., 2011). As for the nature of the emotion change process, how does affect labeling reduce emotional intensity? Dynamic causal modeling analyses identified a role for Broca's area in reducing amygdala activity during affect labeling, but the inhibitory pathway from the VLPFC to the amygdala was much stronger (Torrisi, Lieberman, Bookheimer, & Altshuler, 2013). Given that we know the VLPFC is involved in variety of inhibition and selection processes, these data suggest that affect labeling may engage domain-general inhibitory processes, which dampen amygdala activity and affective responding.

Emotional Stroop and Emotional Go/No Go. Inhibition is the underlying factor driving emotional changes in two other instances of implicit-controlled emotion regulation: the emotional Stroop and go/no-go tasks. In the emotional Stroop, participants are presented with emotional faces overlaid with emotion words (e.g., happy, fear), and their task is to respond to the facial emotion expression while ignoring the word (Etkin, Egner, Peraza, Kandel, & Hirsch, 2006). The process of resolving the conflict generated when the word is incongruent with the facial expression engages cognitive control processes, which carry over to the next trial and facilitate inhibition if the subsequent trial is also incongruent (I-I). Incongruent–incongruent, relative to incongruent trials that are followed by congruent trials (incongruent–congruent), are associated with lower amygdala activity and increased rostral ACC activity (Etkin et al., 2006). A similar effect is thought to occur in the emotional go/no go, in which participants' task is to respond to certain face targets (e.g., females) and withhold a response to others. For no/go trials in which the facial expression was negative, greater activity was observed in the right inferior PFC, DLPFC, and ACC along with decreased activity in amygdala and ventral striatum (Berkman, Burlund, & Lieberman, 2009). Critically, psychophysiological interaction (PPI) analyses demonstrated that the right inferior PFC increases were related to the dampened amygdala activity during no-go trials but not during go trials, which supports the inhibitory spillover hypothesis. Although the focus of the go/no-go and Stroop tasks is not on affect, engaging inhibitory control systems decreases recruitment of the amygdala. Two limitations of the emotional Stroop and emotional go/no-go studies is that they did not assess experience of emotion, only brain activity, and they tested responses to face stimuli, which likely evoke emotion perception processes to a greater degree than they engender affective responding. Despite these limitations, the data suggest that engagement of inhibition decreases affective responding, specifically in the amygdala.

Reversal Learning. Reversal learning describes the process of updating the affective value of a stimulus based on unexpected changes in the outcomes associated with it. For instance, during reversal learning, an organism learns that one stimulus of a pair is initially associated with reward. During reversal, this association flips such that the previously nonrewarded stimulus is now the one linked to reward. Initial accounts of reversal learning suggested that it involves updating the affective values associated with each member of the stimulus pair—a function associated with the VMPFC. This view was supported by animal (for review, see Murray et al., 2007) and human (Fellows, 2011; Fellows & Farah, 2003; Hornak et al., 2004; Rolls, Hornak, Wade, & McGrath, 1994) studies showing reversal deficits after VMPFC lesions as well as the occasional functional magnetic resonance imaging study showing reversal-related VMPFC activity (e.g., Schiller et al., 2008). This view has been challenged, however, by studies in nonhuman primate studies suggesting that VMPFC lesions may damage not just gray matter but fibers of passage connecting the VLPFC to the amygdala and/or striatum. This idea suggests that the VLPFC regions important for the deliberate selection of context-appropriate and inhibition of context-inappropriate responses are what are critical for reversal learning. In keeping with this view are nonhuman primate studies, showing that reversal learning is intact following excitoxic lesions of the VMPFC that damage gray matter but leave passing fibers intact (cf. Rygula, Walker, Clarke, Robbins, & Roberts, 2010); and the majority of human functional magnetic resonance imaging studies of reversal (Cools, Clark, Owen, & Robbins, 2002; Hampshire, Chaudhry, Owen, & Roberts, 2012; Mitchell, Rhodes, Pine, & Blair, 2008; Remijnse, Nielen, Uylings, & Veltman, 2005), showing that the VLPFC is recruited more consistently than the VMPFC during reversal learning.

Automatic Goal Pursuit. When emotion regulation goals are activated outside of awareness, emotions are modulated nonconsciously by a mechanism called automatic goal pursuit (Williams, Bargh, Nocera, & Gray, 2009). Automatic goal pursuit describes the process by which an implicit goal affects higher cognitive processes and behavior in the absence of conscious awareness of the goal (Dijksterhuis & Aarts, 2010). Goals can be activated outside of awareness by priming. Goal priming in the lab typically involves embedding words relevant to the goal in a simple task, such as a word search or scrambled sentences paradigm. Automatic goal pursuit has been studied for a variety of goals including achievement, but only recently has it been studied in the context of emotion regulation. In these studies, emotional control or reappraisal was primed, often by including words related to these concepts in scrambled sentences that participants were tasked with unscrambling (Srull & Wyer, 1979). For example, this scrambled sentence, "drinking restrains she wine from," primes emotional control with the word "restrains" (Mauss, Cook, & Gross, 2007). Emotion control and reappraisal primes, respectively, led to reductions in anger experienced during an anger provocation task (Mauss et al., 2007) and lower physiological reactivity during an anxiety-eliciting task (Williams et al., 2009). Interestingly, reappraisal priming was most effective in individuals who did not habitually use reappraisal in their daily lives (Mauss et al., 2007), suggesting unconsciously triggered regulation operates via a different mechanism from explicit regulation. No neuroimaging studies have been conducted on primed emotion regulation, and future work is needed to identify the neural mechanism of these effects.

Core Neural Systems

Although research to date suggests that implicit-controlled regulation strategies engage combinations of the neural systems observed in explicit-controlled and implicit-automatic regulation, our knowledge concerning hybrid strategies is quite uneven. For example, implicit-controlled regulation has been studied largely using imaging tasks that involve selective attention, which relies on brain regions involved in cognitive control processes, including lateral PFC regions critical for holding in mind goals or selecting/inhibiting responses and posterior medial PFC/dACC regions that monitor performance and evaluate the need for cognitive control (Miller & Cohen, 2001; Wager & Smith, 2003; Wager et al., 2004). It is likely that these brain regions map onto the controlled process of emotion change rather than onto the implicit nature of the goal. Next we summarize what we know about the systems involved in implicit-controlled regulation, but future work is needed to understand the neural systems that support implicit regulation goals.

Lateral PFC and Cingulate Cortex

The lateral PFC and/or the cingulate cortex, as well as modulation of affect-triggering regions like the amygdala, are associated relatively consistently with Stroop tasks (e.g., Buhle, Wager, & Smith, 2010; Ochsner & Gross, 2005). Similarly, greater activity in right inferior PFC, DLPFC, and ACC has been observed on go trials of go/no-go tasks, which may be accompanied by decreased activity and/or altered connectivity with affect-triggering regions, such as the amygdala and ventral striatum (Berkman et al., 2009; Casey et al., 2011; Eigsti et al., 2006; Hare, Tottenham, Davidson, Glover, & Casey, 2005; Hare et al., 2008). The VLPFC, which is important for the deliberate selection of context-appropriate and inhibition of context-inappropriate responses, also has been implicated in reversal learning (Cools et al., 2002; Hampshire et al., 2012; Mitchell et al., 2008; Remijnse et al., 2005; Rygula et al., 2010). However, compare other studies that suggest the VMPFC is important (Fellows, 2011; Fellows & Farah, 2003; Hornak et al., 2004; Murray et al., 2007; Rolls et al., 1994; Schiller, Levy, Niv, Ledoux, & Phelps, 2008).

Explicit-Automatic Regulation

The fourth class of emotion regulation involves mechanisms opposite from those of implicit-controlled regulation. Like explicit-controlled regulation, explicit-automatic regulation requires the presence of an explicit regulatory goal or an expectation that our emotions will change. Unlike explicit-controlled regulation, emotion change occurs relatively automatically. Currently, we have classified one strategy as explicit-automatic regulation: placebo effects.

Prototypical Exemplar: Placebo Effects

Placebo effects arise from the belief that a fake treatment actually alters one's response to a stimulus. In neuroscience research, placebo effects for pain have received the most attention. Numerous studies have shown that consciously holding the belief that a placebo treatment (e.g., a cream) is effective reduces the experience of pain (e.g., unpleasant heat on the forearm). In general, placebo treatment is accompanied by increased engagement of several prefrontal regions, the ventral striatum, and the periaqueductal gray (PAG), and this activity is thought to support maintenance of context information and placebo-related expectations/appraisals (Wager & Atlas, 2015). Placebo treatment leads to reductions in activity in pain-sensitive brain regions (Wager & Atlas, 2015). Placebo beliefs can regulate other types of affective responses, including the experience of disgust and associated insula activity

(Schienle, Übel, Schöngßüner, Ille, & Scharmülle, 2014).

Although expectations of pain relief are held explicitly and engage cortical brain regions, change in affective experience is thought to be relatively automatic. Initial support for an automatic change process is suggested by a behavioral study showing that the effect of placebo-induced analgesia on pain was not reduced by the addition of a concurrent working memory task (Buhle, Stevens, Friedman, & Wager, 2012). The fact that placebo effects persisted suggests that they do not require conscious appraisals/reappraisals of the pain-eliciting stimulus. Reappraisal, in contrast, is thought to depend on explicit reevaluations of stimuli. Recently a study directly compared the neural systems involved in placebo and reappraisal and observed that although both strategies successfully lowered ratings of experienced pain, placebo did so by altering the neural representation of pain, whereas reappraisal altered connectivity between pain-sensitive brain regions and those implicated in reward (Woo et al., 2015).

Core Neural Systems

Compared to explicit and implicit emotion regulation, the neural systems supporting explicit-automatic regulation are not clearly defined. What we do know comes primarily from studies of placebo effects on pain.

DLPFC and VMPFC. Placebo treatment for pain reliably engages the DLPFC and VMPFC, most likely to generate and maintain appraisals and expectations about the effects of the placebo treatment (Wager & Atlas, 2015). The ventral striatum and PAG also typically show enhanced recruitment in the context of placebo treatments. As with other regulation strategies, placebo effects are associated with decreased activity in affective brain regions—here, regions sensitive to pain, including the medial thalamus, anterior insula, and dACC, PAG, and secondary somatosensory cortex-dorsal posterior insula (Wager & Atlas, 2015). Future work with other strategies is needed to understand which brain systems generally support explicit-automatic regulation and which are particular to the kind of affective experience that is being regulated (e.g., pain, disgust, etc.).

ADVANTAGES AND FUTURE DIRECTIONS OF THE FRAMEWORK

The goal of this review was to put forth an organizational framework for understanding different types of emotion regulation in terms of their psychological processes and neural systems. This new two-dimensional framework offers several key benefits. It organizes the existing literature under a more comprehensive scheme that can account for the full range of existing emotion regulation techniques. This framework may improve communication about emotion regulation by standardizing the terms used to describe and categorize emotion regulation work. Finally, we hope that systematically mapping the emotion regulation space will foster new research to address the critical gaps in our knowledge. Having summarized the behavioral and neural evidence for this framework, in this final section, we turn to what we know, what we do not know, and how to integrate and translate this knowledge across different domains of research.

Advantages of the Framework

Understanding emotion regulation strategies by defining them in terms of the dimensions of goal and process is advantageous in at least three ways.

First, it clearly dissociates implicit and explicit forms of regulation at the levels of psychological processes and underlying brain systems. Previously suggested schemes have not cleanly distinguished implicit and explicit emotion regulation, typically because they defined regulation strategies on only one dimension. Although there is no "implicit system" or "explicit system" in the brain, there are brain regions and circuits that are preferentially recruited for each form of regulation. A key neural distinction is revealed by the roles of the lateral PFC, in particular the DLPFC and the VMPFC. Emotion regulation and general "cold" cognitive control research has demonstrated the importance of the DLPFC in representing intentional goals (Miller & Cohen, 2001). In contrast, the representation of implicit goals and context-specific affective value has been linked consistently to the VMPFC (Roy et al., 2012). Anatomical connections between the DLPFC and subcortical brain regions involved in affect, motivation, and learning are sparse, suggesting that the DLPFC plays a critical but indirect role in modulating responding in subcortical brain areas. Consistent with this finding, mediation analysis has demonstrated the presence of a functional DLPFC–ventral striatum pathway in the regulation of craving for appetitive stimuli (Kober et al., 2010b). Interestingly, a recent meta-analysis of studies of reappraisal did not identify the VMPFC as a key region involved in reappraisal (Buhle et al., 2014), furthering the notion that the neural mechanisms of implicit and explicit regulation are dissociated.

Second, moving beyond the implicit-explicit distinction, this framework allows us to parse the extant neuroimaging data more finely and reliably to determine which brain regions/circuits are recruited based on the type of regulatory goal and which are recruited based on the type of change process. This level of neural specificity is critical for characterizing the mechanisms by which regulation occurs. Studying the two hybrid regulation strategies is essential for mapping out this distinction, because they allow us to disentangle the behavioral and neural mechanisms underlying the goal and the change process. As implicit and explicit regulation differ on both the type of goal and the change process, comparisons of implicit and explicit regulation cannot reveal whether differences are due to goal or change process. Although far fewer neuroscience investigations have probed hybrid regulation strategies, we currently know that these forms of regulation involve combinations of the cognitive control and value brain regions that support explicit and implicit regulation.

Third, although the focus of the framework is to clearly distinguish implicit and explicit classes of regulation, it also provides a means for conceptualizing variability in the way in which any individual strategy can be implemented in terms of shifts along the goal and process axes. (See Figure 15.1.) Critically, we believe that individual behavioral strategies are defined by their combination of goal and process. If a given strategy—for example, reappraisal—were implemented in a way that moved it to a different location along the process axis—for example, from engaging many top-down processes to none—we might think of it as a different strategy. For example, after some degree of reappraisal training and/or practice, one may appraise a stimulus as less negative without engaging top-down control processes, and we would think of this as spontaneous reappraisal (Denny & Ochsner, 2014). Similarly, if a strategy moved too far on the goal axis, we might consider it a different strategy. There may be instances where one rethinks the meaning of a situation (reappraises) *without* having the conscious intention to change emotion (no goal); and as reviewed in the implicit-controlled section, the concept of reappraisal can be primed (implicit goal). Would we still call these behaviors reappraisal? We may not yet have specific terms for these variants of regulation strategies, but the framework provides a structure for describing them and making predictions about their neural systems. These examples illustrate that we think of strategies as existing in a cloud that includes the canonical locations depicted in Figure 15.1. We do not yet have enough data to know how far strategies can move from their canonical locations before we think of them as new strategies. However, we have identified a major factor that pushes strategies around in the space, at least strategies that depend on top-down control (explicit- and implicit-controlled regulation): practice/training. Practice likely will reduce the engagement of top-down control processes and potentially also shift the goal from explicit to more implicit. Importantly, the framework offers a means for conceptualizing the various ways in which a given strategy may be instantiated, and changes in the goal and process will be associated with shifts in the underlying neural systems. Thinking about regulatory strategies in this way highlights the need for future work to systematically vary the degree of awareness of regulatory goals and identify factors determining if and when the implementation of a strategy can rely less on top-down control processes.

Future Directions

We hope this framework will be useful for organizing our understanding of emotion regulation, but the day is still young for research in this domain and numerous exciting open questions remain to be addressed in future work. In this section we use the framework to guide discussion of three different directions for future work: further unpacking the nature of the underlying neural systems, understanding the temporal dynamics of regulation, and elucidating factors that influence regulation success, including various kinds of normative and clinically relevant individual differences.

Neural Mechanisms Supporting Explicit and Implicit Regulation

Thus far, we have identified key neural systems supporting explicit and implicit regulation, but additional studies are needed to better describe the critical neural pathways involved in each form of regulation. For instance, we need to delineate the brain systems that are associated with the type of regulatory goal versus the type of change process. Because the majority of extant emotion regulation studies have focused on the change process, we know little about the neural systems that represent emotion regulation goals. We expect that different neural circuitry will be involved in the representation of intentional versus implicit or chronic goals. These distinctions may map onto the lateral PFC and the VMPFC, respectively. The neural systems that support hybrid forms of regulation are less well understood. A goal of specifying this two-dimensional framework is to provide the scaffolding needed to design studies that probe the brain regions critical for the hybrid regulation strategies.

This work could focus on forms of emotion regulation that can be represented in different locations within the framework's two-dimensional space, such that the nature of the goals or change process may vary as a function of how the strategy is implemented. Future work could capitalize on this feature for specific strategies and manipulate it directly. For instance, variants of reappraisal and attentional control strategies can move up and down along the goal dimension, depending on whether the individual possesses a conscious goal to change emotional responding. Indeed, we can change how we think about a stimulus/situation without the intention to change our emotions, just as we might change what we attend to, without regard for the affective consequences of those shifts in attention. Comparing and contrasting the neural systems underlying such variants of reappraisal and attentional control could help clarify the neural systems associated with regulatory goals versus change processes.

Temporal Dynamics of Emotion Regulation

Another set of questions concerns the temporal dynamics of the different forms of regulation. We have identified four major questions regarding temporal dynamics. First, which stage of the emotion generation process is targeted by each form of regulation/specific strategy? These relationships have been clearly spelled out for explicit regulation strategies by Gross (1998a, 1998b) in the emotion regulation process model. We posit that implicit strategies can be mapped roughly onto the same process model, but further examination is needed.

Second, how quickly are regulatory effects observed—that is, how long does it take for regulation to induce emotion changes? For instance, some emotion regulation strategies produce changes in emotion experience right away, such as reappraisal in which decreases in negative emotion are observed on each trial. Other strategies, like extinction, might take numerous trials before significant shifts in affect manifest. Similarly, different forms of emotion regulation may vary as to the length of the delay between the time that the regulatory goal is formed/activated and the time of emotion change. For some strategies, such as reappraisal and distraction, the goal and process stick together, but in others, the time of forming the goal may be temporally extended from the time of change process. With implementation intentions, for instance, the goal to change emotional responding can be set well in advance of encountering the critical affective stimulus. Going forward, it will be important to address differences in these temporal effects.

Third, how long do regulatory effects last? A few studies have addressed this question for specific strategies, but little systematic work has been conducted. For example, a handful of studies suggest that the effects of reappraisal linger. In one event-related potential study, participants viewed negative pictures with either neutral or negative descriptions akin to reappraisals or natural responses and then they viewed the pictures alone 30 minutes later. Pictures previously presented with negative descriptions were rated more negatively and elicited greater neural response in the occipital lobe (assessed with electroencephalogram (MacNamara, Ochsner, & Hajcak, 2011), suggesting that the descriptions had an enduring effect on both experience and brain activity. A second study found that training in cognitive restructuring, a cognitive change strategy akin to reappraisal, was still in effective 24 hours later and led to reduced self-reports of fear and physiological measures of arousal when faced with cues associated with shock (Shurick et al., 2012). And more recently, a pair of imaging studies suggested that the effects of a single reappraisal on diminishing amygdala responses may last 20 minutes (Silvers, Shu, Hubbard, Weber, & Ochsner, 2014a), but if one wanted that diminishment to last up to 1 week, repeatedly reappraising responses to that stimulus four times may be necessary (Denny, Inhoff, Zerubavel, Davachi, & Ochsner, 2015)

A final temporal question is: Does practice or repeated use of a strategy lead to greater effectiveness and an alteration of the underlying neural systems? This question may be particularly important to study explicit-controlled strategies, where practice could make implementation of a goal and/or the change processes more automatic. Little work has investigated this issue, however. One study found that practicing reappraisal over four separate sessions, each several days apart, led to larger reductions in negative affect with time, with biggest reductions seen for reappraisal using a distancing strategy (Denny & Ochsner, 2013). The brain regions recruited also might shift with practice, perhaps from depending on PFC control regions to being supported by VMPFC regions involved in affective updating or even "neutralizing" the affective representation of the stimulus so that engagement of control systems is no longer needed. In keeping with this idea, a recent study found that after reappraising individual stimuli four times, amygdala responses to those stimuli continued to be down-regulated a week later, without the need for PFC control regions to be engaged (Denny et al., 2015). Future research is needed to evaluate such practice-related changes in behavior and brain systems over time.

Factors that Relate to Emotion Regulation Success

Whether a given instance of regulation is successful or unsuccessful may depend on an interaction of person level (e.g., age, clinical status, personality, etc.), situational/stimulus (e.g., what specific emotion triggers are present?) and strategy (e.g., what strategy or strategies is one attempting to employ?) variables (Doré, Silvers, & Ochsner, 2016).

For instance, does choosing the type of regulation strategy—as opposed to being instructed to use a specific strategy—make regulation more effective? We know that participants choose to use distraction more often for high-intensity negative stimuli and reappraisal for low-intensity negative stimuli (Sheppes, Scheibe, Suri, & Gross, 2011), and we know that choice, in general, can enhance motivation and improve behavioral performance (Leotti, Iyengar, & Ochsner, 2010), but we do not yet know the consequences for regulatory success of the choice to regulate per se. Similarly, we know that other factors, such as stress, may impair PFC function in general (Arnsten, 1998), but little is known about how such factors may hinder regulation success. One study has shown that stress can impair explicit regulation (Raio, Orederu, Palazzoli, Shurick, & Phelps, 2013), but little is known about how stress may impact other explicit or implicit strategies.

Perhaps the biggest factor impact regulatory success may be individual differences in emotional reactivity or regulation ability arising from differences in age or clinical disorders. (See Chapter 18 in this volume.) Knowledge of how underlying brain systems change as a function of the stage of life span development or clinical status could be used to determine which kinds of regulation are likely to be the most beneficial for specific individuals. To date, this question has received the most attention in the context of one explicit-controlled strategy—reappraisal—and we close the chapter with a discussion of this work.

Reappraisal Across the Life Span. Although there has been agreement that adolescence is characterized by heightened emotional responsivity, until recently it was unclear whether that should be attributed to increased emotional reactivity, reduced regulation ability, or both. Recent work examining the explicit emotion regulation strategy of reappraisal suggests that the answer to this question depends on the valence of the stimulus in question. On one hand, for aversive stimuli, self-reported emotional reactivity to negative images is relatively constant from childhood through adolescence, but the ability to decrease negative experience via reappraisal improves linearly with age (McRae et al., 2012; Silvers et al., 2012, McRae 2014a). On the other hand, self-reported craving for appetitive food stimuli declines linearly with age while children as young as 6 can use a simple form of reappraisal to successfully down-regulate these cravings (Silvers et al., 2014b). This complementary pattern—age-related changes are seen in regulation for aversive stimuli and in reactivity for appetitive stimuli—highlights that whether children and adolescents, or any population, are able to regulate depends on an interaction of person level, situational/stimulus, and strategy variables.

With respect to older adults, something similar may be at play. Relatively few studies have examined explicit regulation in older adults, but those that have suggest that whether impairments are shown depends on the kinds of strategy involved. For example, a study examining reappraisal of negative images in older and younger adults found that although the groups relied on similar prefrontal brain systems for reappraisal, older adults recruited left VLPFC regions to a lesser extent than younger adults, and individual differences in general cognitive ability—which generally were diminished in older adults—were related to the degree that reappraisal decreased amygdala activity (Winecoff, Labar, Madden, Cabeza, & Huettel, 2011). This study asked participants to use a variant of reappraisal that involves minimizing the negativity of events (e.g., "It's not so bad . . . . Things will be fine . . . . Don't worry . . . ."), a reappraisal tactic that may be more difficult for older adults to use because it depends on cognitive control systems that decline with age but, perhaps more important, does not align with their chronic tendencies to focus on and elaborate positive information (Charles & Carstensen, 2010). Interestingly, behavioral data suggest that when using a reappraisal tactic that capitalizes on this chronic tendency—one that involves finding positive meaning in unpleasant events—older adults may be fine at reappraisal (Shiota & Levenson, 2009). Also in keeping with the idea that impairments depend on the fit of the strategy with the processing capacities and chronic goals of a given group, older adults may be fine at using distraction—an attentional control strategy that depends less on impaired cognitive control systems than does reappraisal—to down-regulate negative emotion (Tucker, Feuerstein, Mende-Siedlecki, Ochsner, & Stern, 2012).

Together, these examples highlight the potential of emotion regulation investigations to inform our understanding of the neural changes occurring across the life span. In turn, knowledge of these changes can be used to determine which forms of regulation may be most beneficial for a given group (Doré et al., 2016)

Reappraisal in Clinical Populations. Many clinical disorders are characterized by emotional disturbances that may stem from deficits in regulation ability, heightened emotional reactivity, or both. These patterns may be disorder specific. For example, individuals suffering from major depressive disorder may fail to recruit a VLPFC-to-amygdala pathway that healthy control participants engage during reappraisal of negative stimuli (Johnstone, Van Reekum, Urry, Kalin, & Davidson, 2007) and may be unable to consistently engage PFC-based control systems that up-regulate ventral striatal activity and maintain positive emotion (Heller et al., 2009). In contrast, individuals with borderline personality disorder may be unable to down-regulate heightened amygdala responses to aversive social stimuli (Silvers et al., 2016), not only because they fail to engage PFC-based control systems effectively but because cingulate regions are not effectively signaling the need for control to come online in the first place (Koenigsberg et al., 2009). Explicit regulation has been investigated in substance-using populations as well, where studies have shown that explicit reappraisal can be used to dampen craving for cigarettes in smokers, and these decreases are related to a functional pathway between the DLPFC and the ventral striatum). One implication of these data is that clinical interventions, such as cognitive behavioral therapy, which rely on many of the same cognitive processes as reappraisal, may depend on the neural systems that support explicit regulation. In some cases, therapeutic interventions may improve their function, but in others, treatments may be contraindicated to the extent that that group cannot effectively engage the processes of interest. In such cases, other types of regulation may be more appropriate, depending on the disorder. We can capitalize on the overlap between emotion regulation strategies and clinical treatments and use neuroimaging investigations of emotion regulation to shed light on the neural mechanisms of clinical treatments.

CONCLUSION

Understanding emotion regulation requires considering its underlying psychological processes, behavioral strategies, and neural systems. In this chapter, we offered an organizational framework that uses these three levels to describe and differentiate four classes of emotion regulation: explicit controlled, implicit automatic, implicit controlled, and explicit automatic. The key feature of the framework is that it describes the psychological processes involved in regulation using two dimensions: the nature of the goal to regulate (implicit or explicit) and the nature of the emotion change process (more automatic or more controlled). Thinking of emotion regulation strategies in this way allows us to understand how neural changes that characterize clinical disorders or accompany different stages of the life span may relate to impaired emotion regulation. We can harness this knowledge to determine which classes of strategies may be most effective for a given population or age group.

REFERENCES

Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (Macaca fascicularis). *Journal of Comparative Neurology*, *230*, 465–496.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience*, *23*, 5627–5633.

Arnsten, A. F. (1998). The biology of being frazzled. *Science*, *280*, 1711–1712.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J. & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *Journal of Neuroscience*, *27*, 3743–3752.

Aron, A. R., Robbins, T. W. & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*, 170–177.

Badre, D. & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, *41*, 473–487.

Bargh, J. A., & Williams, L. E. (2007). The non-conscious regulation of emotion. *Handbook of emotion regulation*, *1*, 429.

Beer, J. S., Heerey, E. A., Keltner, D., Scabini, D. & Knight, R. T. (2003). The regulatory function of self-conscious emotion: insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology*, *85*, 594.

Berkman, E. T., Burklund, L., & Lieberman, M. D. (2009). Inhibitory spillover: intentional motor inhibition produces incidental limbic inhibition via right inferior frontal cortex. *NeuroImage*, *47*, 705–712.

Berkman, E. T., & Lieberman, M. D. (2009). Using neuroscience to broaden emotion regulation: theoretical and methodological considerations. *Social and Personality Psychology Compass*, *3*, 475–493.

Blair, K. S., Smith, B. W., Mitchell, D. G., Morton, J., Vythilingam, M., Pessoa, L., . . . Blair, R. J. (2007). Modulation of emotion by cognition and cognition by emotion. *NeuroImage*, *35*, 430–440.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*, 539–546.

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., . . . Ochsner, K. N. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*, *24*(11), 2981–2990.

Buhle, J. T., Stevens, B. L., Friedman, J. J., & Wager, T. D. (2012). Distraction and placebo: two separate routes to pain control. *Psychological Science*, *23*, 246–253.

Buhle, J., Wager, T., & Smith, E. (2010). Using the Stroop task to study emotion regulation. In R. Hassin, K. N. Ochsner, & Y. Trope (Eds.), *Self Control in Society, Mind, and Brain*, (pp. 93–113). New York, NY: Oxford University Press.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. *Handbook of Emotions* (Vol. 2, pp. 173–191).

Camille, N., Griffiths, C. A., Vo, K., Fellows, L. K., & Kable, J. W. (2011a). Ventromedial frontal lobe damage disrupts value maximization in humans. *Journal of Neuroscience*, *31*, 7527–7532.

Camille, N., Tsuchida, A., & Fellows, L. K. (2011b). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *Journal of Neuroscience*, *31*, 15048–15052.

Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 367–379.

Casey, B. J., Somerville, L. H., Gotlib, I. H., Ayduk, O., Franklin, N. T., Askren, M. K., . . . & Shoda, Y. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 14998–5003.

Cavada, C., Tejedor, J., Cruz-Rizzolo, R. J., & Reinoso-Suárez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cerebral Cortex*, *10*, 220–242.

Charles, S., & Carstensen, L. L. (2010). Social and emotional aging. *Annual Review of Psychology*, *61*, 383.

Chib, V. S., Rangel, A., Shimojo, S. & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, *29*, 12315–12320.

Cohen, J. R., & Lieberman, M. D. (2010). The common neural basis of exerting self-control in multiple domains. In R. Hassin, K. Ochsner, & Y. Trope (Eds.), *Self Control in Society, Mind, and Brain* (pp.141–162). New York, NY: Oxford University Press.

Cools, R., Clark, L., Owen, A. M. & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 4563–4567.

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. 2008. Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews*, *58*, 57–70.

Craig, A. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*, 59–70.

Critchley, H. [D.], Daly, E., Phillips, M., Brammer, M., Bullmore, E., Williams, S., . . . Murphy, D. (2000). Explicit and implicit neural mechanisms for processing of social information from facial expressions: a functional magnetic resonance imaging study. *Human Brain Mapping*, *9*, 93–105.

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*, 189–195.

Cunningham, W. A., & Brosch, T. 2012. Motivational salience: amygdala tuning from traits, needs, values, and goals. *Current Directions in Psychological Science*, *21*, 54–59.

Curtis, C. E., & D'Esposito, M. 2003. Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, *7*, 415–423.

Custers, R., & Aarts, H. 2010. The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science*, *329*, 47–50.

Davis, M., & Whalen, P. J. 2001. The amygdala: vigilance and emotion. *Molecular Psychiatry*, *6*, 13–34.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*, 70–88.

Delgado, M. R., Gillis, M. M., & Phelps, E. A. (2008). Regulating the expectation of reward via cognitive strategies. *Nature Neuroscience*, *11*, 880–881.

Delgado, M. R., Olsson, A. & Phelps, E. A. (2006). Extending animal models of fear conditioning to humans. *Biological Psychology*, *73*, 39–48.

Denny, B. T., Inhoff, M. C., Zerubavel, N., Davachi, L. & Ochsner, K. N. (2015). Getting over it: long-lasting effects of emotion regulation on amygdala response. *Psychological Science*, *26*(9),1377–88. doi:10.1177/0956797615578863.

Denny, B. T., & Ochsner, K. N. (2014). Behavioral effects of longitudinal training in cognitive reappraisal. *Emotion*, *14*, 425–433.

Diekhof, E. K., Geier, K., Falkai, P., & Gruber, O. (2011). Fear is only as deep as the mind allows: a coordinate-based meta-analysis of neuroimaging studies on the regulation of negative affect. *NeuroImage*, *58*, 275–285.

Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un) consciousness. *Annual Review of Psychology*, *61*, 467–490.

Doré, B. P., Silvers, J. A., & Ochsner, K. N. 2016. Toward a personalized science of emotion regulation. *Social and Personality Psychology Compass*, *10*, 171–187.

Eigsti, I. M., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., . . . Casey, B. J. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychological Science*, *17*, 478–84.

Elliott, R., Agnew, Z., & Deakin, J. (2008). Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans. *European Journal of Neuroscience*, *27*, 2213–2218.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*, 871–882.

Fellows, L. K. (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Annals of the New York Academy of Sciences*, *1239*, 51–58.

Fellows, L. K., & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain*, *126*, 1830–1837.

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex*, *17*, 2669–2674.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*, *301*, 1104–1107.

Green, M. J., & Malhi, G. S. (2006). Neural mechanisms of the cognitive control of emotion. *Acta Neuropsychiatrica*, *18*, 144–153.

Gross, J. J. (1998a). Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*, 224–237.

Gross, J. J. (1998b). The emerging field of emotion regulation: an integrative review. *Review of General Psychology*, *2*, 271–299.

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: one or two depends on your point of view. *Emotion Review*, *3*, 8–16.

Haber, S. N., & Knutson, B. 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*, 4–26.

Hampshire, A., Chaudhry, A. M., Owen, A. M., & Roberts, A. C. (2012). Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *NeuroImage*, *59*, 4102–4112.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, *324*, 646–648.

Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H., & Casey, B. J. (2005). Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry*, *57*, 624–632.

Hare, T. A., Tottenham, N., Galvan, A., Voss, H. U., Glover, G. H., & Casey, B. J. (2008). Biological substrates of emotional reactivity and regulation in adolescence during an emotional go-no go task. *Biological Psychiatry*, *63*, 927–934.

Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport*, *11*, 43–48.

Hartley, C. A., & Phelps, E. A. 2009. Changing fear: the neurocircuitry of emotion regulation. *Neuropsychopharmacology*, *35*, 136–146.

Hassin, R. R. (2013). Yes it can: on the functional abilities of the human unconscious. *Perspectives on Psychological Science*, *8*, 195–207.

Hassin, R. R., Aarts, H., Eitam, B., Custers, R., & Kleiman, T. (2009). Non-conscious goal pursuit and the effortful control of behavior. In E. Morsella & J. A. Bargh (Eds.), *Oxford Handbook of Human Action*, pp. 549–566. New York, NY: Oxford University Press.

Hassin, R. R., Ochsner, K., & Trope, Y. (2010). *Self Control in Society, Mind, and Brain*. New York, NY: Oxford University Press.

Heberlein, A. S., Padon, A. A., Gillihan, S. J., Farah, M. J., & Fellows, L. K. (2008). Ventromedial frontal lobe plays a critical role in facial emotion recognition. *Journal of Cognitive Neuroscience*, *20*, 721–33.

Heller, A. S., Johnstone, T., Shackman, A. J., Light, S. N., Peterson, M. J., Kolden, . . . Davidson, R. J. (2009). Reduced capacity to sustain positive emotion in major depression reflects diminished maintenance of fronto-striatal brain activation. *Proceedings of the National Academy of Sciences*, *106*, 22445–22450.

Holland, P. C., & Gallagher, M. 2004. Amygdala–frontal interactions and reward expectancy. *Current Opinion in Neurobiology*, *14*, 148–155.

Hornak, J., Bramham, J., Rolls, E. T., Morris, R. G., O'Doherty, J., Bullock, P. R., & Polkey, C. E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain*, *126*, 1691–2712.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E., Morris, R., Bullock, P. & Polkey, C. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, *16*, 463–478.

Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*, *34*, 247–61.

Janak, P. H., & Tye, K. M. (2015). From circuits to behaviour in the amygdala. *Nature*, *517*, 284–92.

Johnstone, T., Van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. 2007. Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *Journal of Neuroscience*, *27*, 8877–8884.

Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A. & Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*, *338*, 953–956.

Kanske, P., Heissler, J., Schönfelder, S., Bongers, A., & Wessa, M. (2010). How to regulate emotion? Neural networks for reappraisal and distraction. *Cerebral Cortex*, *21*, 1379–1388.

Kerfoot, E. C., Agarwal, I., Lee, H. J., & Holland, P. C. (2007). Control of appetitive and aversive taste-reactivity responses by an auditory conditioned stimulus in a devaluation task: a FOS and behavioral analysis. *Learning & Memory*, *14*, 581–589.

Knight, R. T., Staines, W. R., Swick, D., & Chao, L. L. (1999). Prefrontal cortex regulates inhibition and excitation in distributed neural networks. *Acta Psychologica*, *101*, 159–178.

Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage*, *42*, 998–1031.

Kober, H., Kross, E. F., Mischel, W., Hart, C. L., & Ochsner, K. N. (2010a). Regulation of craving by cognitive strategies in cigarette smokers. *Drug and Alcohol Dependence*, *106*, 52–55.

Kober, H., Mende-Siedlecki, P., Kross, E. F., Weber, J., Mischel, W., Hart, C. L., & Ochsner, K. N. (2010b). Prefrontal-striatal pathway underlies cognitive regulation of craving. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14811–14816.

Koenigsberg, H. W., Fan, J., Ochsner, K. N., Liu, X., Guise, K. G., Pizzarello, S., . . . Goodman, M. (2009). Neural correlates of the use of psychological distancing to regulate responses to negative social cues: a study of patients with borderline personality disorder. *Biological Psychiatry*, *66*, 854–863.

Ledoux, J. E. (2012). Evolution of human emotion: a view through fear. *Progress in Brain Research*, *195*, 431.

Leotti, L. A., Iyengar, S. S., & Ochsner, K. N. (2010). Born to choose: the origins and value of the need for control. *Trends in Cognitive Sciences*, *14*, 457–463.

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science*, *18*, 421–428.

Lieberman, M. D., Inagaki, T. K., Tabibnia, G., & Crockett, M. J. (2011). Subjective responses to emotional stimuli during labeling, reappraisal, and distraction. *Emotion*, *11*, 468.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. 2012. The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences*, *35*, 121–143.

MacNamara, A., Ochsner, K. N., & Hajcak, G. 2011. Previously reappraised: the lasting effect of description type on picture-elicited electrocortical activity. *Social Cognitive and Affective Neuroscience*, *6*, 348–358.

Martin, L. N., & Delgado, M. R. (2011). The influence of emotion regulation on decision-making under risk. *Journal of Cognitive Neuroscience*, *23*, 2569–2581.

Mauss, I. B., Cook, C. L., & Gross, J. J. (2007). Automatic emotion regulation during anger provocation. *Journal of Experimental Social Psychology*, *43*, 698–711.

McRae, K., Gross, J. J., Weber, J., Robertson, E. R., Sokol-Hessner, P., Ray, R. D., . . . Ochsner, K. N. (2012). The development of emotion regulation: an fMRI study of cognitive reappraisal in children, adolescents and young adults. *Social Cognitive and Affective Neuroscience*, *7*, 11–22.

McRae, K., Hughes, B., Chopra, S., Gabrieli, J. D., Gross, J. J. & Ochsner, K. N. (2010). The neural bases of distraction and reappraisal. *Journal of Cognitive Neuroscience*, *22*, 248–62.

Milad, M. R., Quinn, B. T., Pitman, R. K., Orr, S. P., Fischl, B., & Rauch, S. L. (2005). Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 10706–10711.

Milad, M. R., & Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature*, *420*, 70–74.

Miller, B. T. & D'Esposito, M. (2005). Searching for "the top" in top-down control. *Neuron*, *48*, 535–538.

Miller, E. K., & Cohen, J. D. 2001. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Milne, E., & Grafman, J. (2001). Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *Journal of Neuroscience*, *21*, 1–6.

Mitchell, D., Rhodes, R., Pine, D., & Blair, R. (2008). The contribution of ventrolateral and dorsolateral prefrontal cortex to response reversal. *Behavioural Brain Research*, *187*, 80–87.

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*, 119–124.

Murray, E. A., O'Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *Journal of Neuroscience*, *27*, 8166–8169.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, *14*, 1215–1229.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*, 242–249.

Ochsner, K. N., & Gross, J. J. (2007). The neural architecture of emotion regulation. In J. J. Gross (Ed.), *Handbook of Emotion Regulation* (pp. 87–109). New York, NY: Guilford Press.

Ochsner, K. N., & Gross, J. J. (2008). Cognitive emotion regulation: insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science*, *17*, 153–158.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage*, *23*, 483–499.

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Science*, *1251*, E1–E24.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology*, *14*, 769–776.

Öngür, D., & Price, J. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, *10*, 206–219.

Panksepp, J. (2003). At the interface of the affective, behavioral, and cognitive neurosciences: decoding the emotional feelings of the brain. *Brain and Cognition*, *52*, 4–14.

Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, *8*, 162–166.

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*, 160–166.

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology*, *57*, 27–53.

Phelps, E. A., Delgado, M. R., Nearing, K. I., & Ledoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron*, *43*, 897–905.

Pickens, C. L., & Holland, P. C. (2004). Conditioning and cognition. *Neuroscience & Biobehavioral Reviews*, *28*, 651–661.

Plassmann, H., O'Doherty, J., & Rangel, A. 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, *27*, 9984–9988.

Price, J. L. (2007). Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Annals of the New York Academy of Sciences*, *1121*, 54–71.

Price, J. L., & Drevets, W. C. (2009). Neurocircuitry of mood disorders. *Neuropsychopharmacology*, *35*, 192–216.

Quirk, G. J., & Mueller, D. (2007). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*, *33*, 56–72.

Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *Journal of Neuroscience*, *20*, 6225–6231.

Raio, C., Orederu, T., Palazzolo, L., Shurick, A., & Phelps, E. (2013). Cognitive emotion regulation fails the stress test. *Proceedings of the National Academy of Sciences*, *110*, 15139–15144.

Remijnse, P. L., Nielen, M. M., Uylings, H. B., & Veltman, D. J. (2005). Neural correlates of a reversal learning task with an affectively neutral baseline: an event-related fMRI study. *NeuroImage*, *26*, 609–618.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. 1994. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery & Psychiatry*, *57*, 1518–1524.

Roy, M., Shohamy, D., & Wager, T. D. 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends in Cognitive Sciences*, *16*, 147–156.

Rygula, R., Walker, S. C., Clarke, H. F., Robbins, T. W., & Roberts, A. C. (2010). Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *Journal of Neuroscience*, *30*, 14552–14559.

Salas, C. E., Gross, J. J., Rafal, R. D., Viñas-Guasch, N., & Turnbull, O. H. (2013). Concrete behaviour and reappraisal deficits after a left frontal stroke: a case study. *Neuropsychological Rehabilitation*, *23*, 467–500.

Salas, C. E., Gross, J. J., & Turnbull, O. H. (2014). Reappraisal generation after acquired brain damage: the role of laterality and cognitive control. *Frontiers in Psychology*, *5*, 242. doi: 10.3389/fpsyg.2014.00242

Scherer, K. R., Schorr, A. E., & Johnstone, T. E. 2001. *Appraisal Processes in Emotion: Theory, Methods, Research*. New York, NY: Oxford University Press.

Schienle, A., Übel, S., Schöngßüner, F., Ille, R., & Scharmüller, W. (2014). Disgust regulation via placebo: an fMRI study. *Social, Cognitive and Affective Neuroscience*, *9*(7), 985–990.

Schiller, D., Levy, I., Niv, Y., Ledoux, J. E. & Phelps, E. A. (2008). From fear to safety and back: reversal of fear in the human brain. *Journal of Neuroscience*, *28*, 11517–11525.

Schmeichel, B. J., Volokhov, R. N., & Demaree, H. A. (2008). Working memory capacity and the self-regulation of emotional expression and experience. *Journal of Personality and Social Psychology*, *95*, 1526.

Schoenbaum, G., Takahashi, Y., Liu, T. Ä., & McDannald, M. A. (2011). Does the orbitofrontal cortex signal value? *Annals of the New York Academy of Sciences*, *1239*, 87–99.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.

Sheppes, G., Scheibe, S., Suri, G., & Gross, J. J. (2011). Emotion-regulation choice. *Psychological Science*, *22*, 1391–1396.

Shidlovski, D., & Hassin, R. R. (2011). When pooping babies become more appealing: the effects of nonconscious goal pursuit on experienced emotions. *Psychological Science*, *22*, 1381–1385.

Shiota, M. N., & Levenson, R. W. (2009). Effects of aging on experimentally instructed detached reappraisal, positive reappraisal, and emotional behavior suppression. *Psychology and Aging*, *24*, 890.

Shurick, A. A., Hamilton, J. R., Harris, L. T., Roy, A. K., Gross, J. J., & Phelps, E. A. (2012). Durable effects of cognitive restructuring on conditioned fear. *Emotion*, *12*, 1393.

Silvers, J. A., Hubbard, A. D., Biggs, E., Shu, J., Fertuck, E., Chaudhury, S., . . . Chesin, M. (2016). Affective lability and difficulties with regulation are differentially associated with amygdala and prefrontal response in women with borderline personality disorder. *Psychiatry Research: Neuroimaging*, *254*, 74–82.

Silvers, J. A., Insel, C., Powers, A., Franz, P., Weber, J., Mischel, W., . . . & Ochsner, K. N. (2014b). Curbing craving behavioral and brain evidence that children regulate craving when instructed to do so but have higher baseline craving than adults. *Psychological Science*, *25*, 1932–1942.

Silvers, J. A., McRae, K., Gabrieli, J. D., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2012). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. *Emotion*, *12*, 1235.

Silvers, J. A., Shu, J., Hubbard, A. D., Weber, J. & Ochsner, K. N. (2014a). Concurrent and lasting effects of emotion regulation on amygdala response in adolescence and young adulthood. *Developmental Science*.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661.

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: some determinants and implications. *Journal of Personality and Social Psychology*, *37*, 1660.

Stjepanovic, D. L. K. (2017). Fear learning. In J. T. Wixted, E. Phelps, & L. Davachi (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 1: Learning & Memory). Hoboken, NJ: Wiley.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. 1998. Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*, 640–656.

Tabibnia, G., Monterosso, J. R., Baicy, K., Aron, A. R., Poldrack, R. A., Chakrapani, S., . . . London, E. D. 2011. Different forms of self-control share a neurocognitive substrate. *Journal of Neuroscience*, *31*, 4805–4810.

Taylor, S. F., Phan, K. L., Decker, L. R., & Liberzon, I. (2003). Subjective rating of emotionally salient stimuli modulates neural activity. *NeuroImage*, *18*, 650–659.

Tice, D. M., & Bratslavsky, E. 2000. Giving in to feel good: the place of emotion regulation in the context of general self-control. *Psychological Inquiry*, *11*, 149–159.

Torrisi, S. J., Lieberman, M. D., Bookheimer, S. Y., & Altshuler, L. L. (2013). Advancing understanding of affect labeling with dynamic causal modeling. *NeuroImage*, *15*(82), 481–488.

Tucker, A. M., Feuerstein, R., Mende-Siedlecki, P., Ochsner, K. N., & Stern, Y. 2012. Double dissociation: circadian off-peak times increase emotional reactivity; aging impairs emotion regulation via reappraisal. *Emotion*, *12*, 869.

Van Dillen, L. F., Heslenfeld, D. J., & Koole, S. L. 2009. Tuning down the emotional brain: an fMRI study of the effects of cognitive load on the processing of affective images. *NeuroImage*, *45*, 1212–1219.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*, *30*, 829–841.

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *Journal of Cognitive Neuroscience*, *22*, 2864–2885.

Wager, T. D., & Atlas, L. Y. (2015). The neuroscience of placebo effects: connecting context, learning and health. *Nature Reviews Neuroscience*, *16*, 403–418.

Wager, T. D., Atlas, L. Y., Lindquist, M. A., Roy, M., Woo, C. W., & Kross, E. 2013. An fMRI-based neurologic signature of physical pain. *New England Journal of Medicine*, *368*, 1388–1397.

Wager, T. D., Jonides, J., & Reading, S. 2004. Neuroimaging studies of shifting attention: a meta-analysis. *NeuroImage*, *22*, 1679–1693.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 255–274.

Webb, T. L., Schweiger Gallo, I., Miles, E., Gollwitzer, P. M., & Sheeran, P. (2012). Effective regulation of affect: an action control perspective on emotion regulation. *European Review of Social Psychology*, *23*, 143–186.

Williams, L. E., Bargh, J. A., Nocera, C. C., & Gray, J. R. (2009). The unconscious regulation of emotion: nonconscious reappraisal goals modulate emotional reactivity. *Emotion*, *9*, 847–854.

Winecoff, A., Labar, K. S., Madden, D. J., Cabeza, R., & Huettel, S. A. (2011). Cognitive and neural contributors to emotion regulation in aging. *Social Cognitive and Affective Neuroscience*, *6*, 165–176.

Winston, J., O'Doherty, J., & Dolan, R. 2003. Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage*, *20*, 84–97.

Woo, C. W., Roy, M., Buhle, J. T., & Wager, T. D. (2015). Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain. *PLoS Biology*, *13*, e1002036.

Zaki, J., Davis, J. I., & Ochsner, K. N. (2012). Overlapping activity in anterior insula during interoception and emotional experience. *NeuroImage*, *62*, 493–499.

CHAPTER 16

Self-Evaluation

JENNIFER S. BEER, ANASTASIA E. RIGNEY, AND JESSICA E. KOSKI

People's intrinsic interest in gaining self-knowledge is reflected in the numerous quizzes that promise to tell them which TV or movie character shares their personality, the joking expression "But enough about me, what do you think about me?," and the billion-dollar self-help industry. This chapter focuses on understanding self-knowledge from a psychological and neurobiological perspective. From a psychological perspective, one unique aspect of self-knowledge is that the perceiver is the same person as the perceived. In contrast to cases where we form impressions of other people without access to their life histories or innermost thoughts and emotions, our self-knowledge is acquired by us about us. In this chapter, we consider the definition of the term "self" and then address a number of research themes that consider how the simultaneous roles of perceive and perceived contribute to self-knowledge. For example, what processes explain how we learn about ourselves, and are they similar or different to the way we learn about other things? Furthermore, how accurate is our process of learning about ourselves? Can we dispassionately assemble information, or do we fudge the data about ourselves to maintain certain self-views? Once self-knowledge has been acquired, is it represented in any special way or in the same way as other-person knowledge? How do these processes relate at the psychological level of analyses and neurobiological levels of analyses? Decades of research suggests that we draw on both internal and external sources of self-information but rarely collect this information in a dispassionate manner. Additionally, psychological and neural research suggest we represent self-knowledge in a particularly rich manner, but the underlying psychological and neural processes that support self-knowledge are much more similar to knowledge about other people than originally thought.

WHAT WE TALK ABOUT WHEN WE TALK ABOUT THE SELF

The word "self" is used in many ways in everyday language, but what are scientists referring to when they use the term "self"? Although minor differences exist, most psychological models of the self-characterize it with internal, external, and socially perceived attributes that are influenced by many variables, such as culture, time, and motivation. The classic definition of self comes from William James (1890/1983). James defines the self as the material, social, and spiritual aspects of the perceived self as well as the perception of these aspects. According to James, the "material" self includes one's physical body as well as possessions and the people you call family. The "social" self refers to the way in which people represent themselves in their own minds as well as how other people view them (i.e., a social representation). In other words, one's social reputation is just as much part of the self as self-perceptions. The "spiritual" self is characterized by internal qualities such as personality, attitudes, and consciousness. Although all three aspects are used to define the self, subsequent research has shed light on the ways in which they cohere and the ways in which they do not. Furthermore, researchers have discovered that a number of factors influence which aspects are especially important to self-knowledge; these factors include temporal construal, culture, and motivation (e.g., Markus & Kitayama, 1991; Markus & Nurius, 1986; Sedikides & Gregg, 2008; Swann, Pelham, & Krull, 1989).

When External, Internal, and Social Representations of Self Align

Plenty of evidence suggests coherence among the external, internal, and social representations of self. One example of how one's internal attributes, external attributes, and social representations can be interrelated is illustrated by the operation of the "beautiful is good" stereotype (e.g., Bhanji & Beer, 2013; Dion, Berscheid, & Walster, 1972; Meier, Robinson, Carter, & Hinsz, 2010). People tend to judge physically attractive individuals as having more desirable personalities; physically attractive people also view themselves as having more desirable personalities. In the seminal research on this effect, participants viewed photographs of three people and formed impressions of their personalities (Dion et al., 1972). The physical attractiveness of the individuals in the photographs was manipulated such that one was unattractive, one had average attractiveness, and one was very attractive. The attractive individual was perceived as more outgoing and sociable than the unattractive or average attractive individual. Further research illustrated how the correspondence between social representations of individuals and their physical attributes may also relate to internal attributes. For example, there is agreement between observer reports that use physical attractiveness cues to judge sociability and self-reports of one's own sociability (Meier et al., 2010). Participants provided self-reports of their personalities and posed for a spontaneous photograph. There was a significant degree of correspondence between self-reports of sociability and observer reports of physical attractiveness.

Additionally, personality can predict a person's clothes, possessions, and content of their social media (Gosling, Ko, Mannarelli, & Morris, 2002). The physical attributes of someone's living space (i.e., external attributes) significantly correlate with self-perceptions of personality (i.e., internal attributes) as well as friends' perceptions of that person's personality (i.e., the social self). For example, in one study, observers made personality ratings after viewing the bedrooms of target individuals. These personality ratings were then considered in relation to self-perceptions and friend perceptions of the target's personality. Significant agreement was found for how much a target tended to be extraverted, agreeable, open to new experiences, conscientious, and neurotic. Furthermore, it was shown that observer-reported and self-reported personality tended to rely on similar physical attributes of the self. For example, targets were viewed as highly conscientious and rated themselves as highly conscientious to the extent that their bedrooms were well lighted, organized, and not cluttered with books or CDs. Similarly, targets were viewed as highly open to new experiences and considered themselves to be highly open to new experiences to the extent that their bedrooms had unique décor and diverse reading material. Taken together, these studies illustrate how external attributes such as possessions or physical appearance relate to internal attributes and affect how the self is represented in the minds of other people.

Differences Among External, Internal, and Social Representations of Self

However, there is not always high correlation among the self's external attributes, internal attributes, and social representations. This principle is illustrated by research examining the ways in which people erroneously make inferences about a target's internal attributes because of their own motivations or on the basis of the objects in their rooms. For example, recent research suggests that stereotypes, including the "beautiful is good" stereotype, stem partly from an observer's motivation. When observers' desire to affiliate with physically attractive people was statistically controlled, many of the relations between a target's physical attractiveness and observer-rated affiliative personality characteristics were weakened or nonexistent (Lemay, Clark, & Greenberg, 2010). These findings suggest that at least some of the "beautiful is good" stereotype comes from observers' projections of their desire to affiliate with physically attractive individuals, which subsequently motivates them to construe those individuals as likely to want to affiliate in return. Additionally, in the bedroom study mentioned earlier, some of the external attributes that were used to judge targets' personality were not related to self-report or friend reports of personality (Gosling et al., 2002). For example, although colorful bedrooms tended to garner high ratings of agreeableness and conscientiousness, colorful bedrooms did not predict self-reports of those traits.

Culture and Self-Definition

One other reason that external attributes, internal attributes, and social representations may not cohere as much as predicted by James's classic definition of the self is that culture may influence the extent to which the external attributes, internal attributes, or social representation is considered to form the core of the self. For example, cultural influences predict how much family members or other social groups are construed as part of the self. One of the largest research programs on cultural influences of self has focused on how much cultures vary in construing the self as independent or interdependent in relation to social groups (Markus & Kitayama, 1991). Independent self-construals emphasize the self's core features as those external and internal attributes that distinguish an individual from other people. People are "acting like themselves" as long as their appearance and actions are being driven by their internal attributes. In contrast, interdependent self-construals emphasize the self's core features as those that connect them to a larger social entity. In cultures where people are taught to construe themselves through an interdependent lens, they are strongly defined by their family and other social groups (considered to be material aspects of the self by James, 1890) as well as how they are represented in the minds of other people (considered to be the social self by James, 1890). In other words, self-expression is not considered to rest on expressing one's unique qualities but more about one's role within the group.

One of the most widely used paradigms to understand cultural influences on the self is the "I am . . ." test (e.g., Bochner, 1994; Bond & Cheung, 1983). Studies using this approach ask participants to complete the sentence "I am . . ." using whatever phrases they deem appropriate. Participants with interdependent self-construals are more likely to use social roles or relationships in their sentence completions. For example, they may complete the sentences by noting that they are a mother, a daughter, and a sister. In contrast, participants whose culture emphasizes independent self-construals are more likely to identify their unique personality characteristics when completing these sentences. For example, they may complete the sentences by noting that they are intelligent, strong, and talented. The research on independent and interdependent construals of self exemplifies the way in which culture may affect the coherence among the material, social, and internal aspects of self as they may receive differential importance in self-definition.

Temporal Construal and the Self

Time also affects coherence of the self. The nature of our inner attributes may change with time; we distinguish among our past, current, and future inner states (e.g., Bartels & Rips, 2010; Markus & Nurius, 1986; Trope & Liberman, 2010; Wilson & Gilbert, 2005). These temporal construals of self are associated with different preferences. For example, people are willing to accept a smaller amount of money today over a larger amount of money that they would receive at a later date. Similarly, people tend to believe that they will choose to drink significantly less of an unpleasant liquid to advance scientific knowledge if the drinking will occur today compared to a future date (e.g., 3 months from now; Pronin, Olivola, & Kennedy, 2008). Our different representations of self across time are one explanation offered for why we make such different choices for ourselves across time. It is theorized that we identify more closely with our present selves, and, therefore, we tend to be more protective (i.e., maximize benefits and minimize costs) of our current selves than our future selves (Parfit, 1984). These studies illustrate that the inner attributes that James used to define the self may differ across time and, therefore, may cohere with external attributes and social representations differentially across time.

Summary

James argued that the self consists of two constituents: the perceived and the perceiver. The self consists of the attributes that are visible to others as well as less tangible internal attributes and the representation of both kinds of attributes in our own minds and the minds of other people (i.e., our social reputation). These aspects of self can cohere, but they do not always. Culture affects how much we favor different aspects when defining the self, and temporal construal affects our perception of the inner attributes of the perceived self. Motivational influences on self-evaluation are discussed later in the chapter.

PROCESSES OF SELF-EVALUATION

What are the ways in which we evaluate our external attributes, internal attributes, and social representations? Several different theories, such as the self-perception theory (Bem, 1967, 1972), the looking-glass self (Cooley, 1902; Mead, 1934), and social comparison (Festinger, 1954; Goethals & Darley, 1977; Kruglanski & Mayseless, 1990; Mussweiler & Rueter, 2003) explain these processes, and all likely capture the myriad avenues that we take to gain self-knowledge. When these theories are considered, the process of getting to know oneself is similar to online shopping. You cannot touch online goods or see inside them before your purchase. So how do you know what the product is like? Do you observe what you can about product, such as its visible attributes and the information provided about the product's features? Do you look to see what other people think of the product? Both processes are helpful, and, when applied to acquiring self-knowledge, circumstances may determine which process takes center stage.

Self-Perception Theory

According to self-perception theory (Bem, 1967, 1972), self-knowledge is gained much in the same way we gain knowledge about other people (or even online products): by observation. We can never access the mind of another person in a firsthand manner. Therefore, one avenue for learning about them is to observe them. Similarly, we may try to understand ourselves by noting our behavior. If we couch this process in Jamesian terms, self-perception theory posits that the perceiver self learns about internal attributes by observing the external aspects of the perceived self. Although this perspective may seem at odds with one's own introspection on self-knowledge acquisition, there is empirical evidence to support this view (Festinger, 1957; Festinger & Carlsmith, 1959; Nisbett & Wilson, 1977). However, it is important to note that this method may not be the primary means of acquiring self-knowledge. People report that their innermost thoughts and emotions are more central to defining themselves than their observations of their behavior (Andersen & Ross, 1984). For example, participants were asked to rate how much someone could learn about them from a sample of their internal thoughts and feelings compared to a sample of their external behaviors. Participants felt both sources of information would be informative but rated the sample of thoughts and feelings as significantly more informative than the sample of external behavior (Andersen & Ross, 1984).

Currently, researchers do not focus as much on debating the relative importance of behavioral observation versus introspection for gaining self-knowledge. Instead, the focus is on understanding the convergence and divergence in the methods used to learn about ourselves versus other people. Do we tend to rely on observation more when trying to understand other people compared to understanding ourselves? There is certainly evidence for differences in the way we think about the best way to try to understand ourselves compared to the best way to understand other people (e.g., Jones & Nisbett, 1971; Pronin et al., 2002). For example, in the Andersen and Ross (1984) study mentioned earlier, participants reported that observation of external behavior would likely be more informative about another person than it would be about themselves.

Yet the manner in which we explain our behavior parallels the way in which we make attributions about other people's behavior (e.g., Knee, Patrick, & Lonsbury, 2003; Malle, 2006; Nisbett & Wilson, 1977; Plaks, Levy, & Dweck, 2009; Taylor & Koivumaki, 1976). Students of classic social psychology might take issue with that claim—people were classically considered to focus more on situational explanations of their own behavior compared to explanations for another person's behavior (Jones & Nisbett, 1971). In other words, if we asked you why you did not give up your seat on the bus for a stranger with a bulky bag, you would likely look for something special about that particular instance to explain your behavior (e.g., you were feeling poorly that day). However, if we asked you why someone else failed to give up their seat to someone in need, you would be more likely to attribute their behavior to something about their disposition (e.g., they are an inconsiderate person). However, a meta-analysis has shown that these differences are not as robust as previously thought (Malle, 2006); people are likely to make dramatically different attributions only in circumstances where information suggests real differences between the self and other people.

Looking-Glass Self

James argued that a crucial part of ourselves is how we are represented in other people's minds, yet self-perception theory explains self-evaluation with a process that operates in the absence of other people. Do other people play a role in the way we come to know ourselves? At least two theories suggest that yes, other people do play an important role in gaining self-knowledge. The first theory, the looking-glass self, posits that people help us learn about ourselves by communicating what they see in us. Other people act as looking glasses, or mirrors, in which we can observe ourselves. Rather than the direct observations that form the core of self-perception theory, the looking-glass self perspective focuses on our observations of what other people see in us. More specifically, people are thought to imagine how they are judged by other people (i.e., reflected appraisals) and then internalize their judgments (e.g., Cooley, 1902; Mead, 1934).

Research has shown that we are more likely to incorporate information about our behavior into our self-view if we believe the behaviors were observed by another person. For example, participants were asked to serve as test cases for graduate students who were being trained in clinical observation (Tice, 1992). Participants were instructed to present themselves as emotionally stable, emotionally unstable, or possessing a task-irrelevant attribute (i.e., athletic). Participants were then randomly assigned to one of two conditions: one condition in which they believed that they were interacting with a graduate student who could see them through a one-way mirror or another condition in which they believed they were being recorded for future training. After participants responded to a series of questions in which they had a chance to portray themselves as instructed, they filled out questionnaires that were presumably unrelated to the first task. However, the questionnaires included a self-assessment of emotional stability. Participants were most likely to rate themselves as possessing their assigned level of emotional stability if they believed they had been watched rather than merely recorded. A follow-up study found that the effects of public versus private portrayals on incorporation into self-views did generalize to other personality characteristics, such as extraversion (Tice, 1992). This study illustrates that situations which evoke thoughts about how you are being perceived by another person can affect your perception of your personality.

Another implication of the looking-glass self is that self-evaluations should be consistent with how other people perceive you and that people should be aware of how they are perceived by others. Consistent with this prediction, correlational research finds a high level of agreement between how people perceives themselves and how they are perceived by others, at least on certain dimensions (e.g., Albright, Kenny, & Malloy, 1988; Ambady & Rosenthal, 1992; Gosling et al., 2002; Marcus & Miller, 2003; Norman & Goldberg, 1966). Studies of students who meet on the first day of class illustrate this body of work. In one study, students on the first day of class were assigned to groups, and a round-robin design was used to collect physical attractiveness ratings (Marcus & Miller, 2003). In other words, each group member rated themselves and all other group members on attractiveness and also rated how they believed their attractiveness was judged by each group member. The participants' attractiveness ratings were correlated with how other people see them. Additionally, participants were aware of how they were seen by others in terms of physical attractiveness. But did these effects generalize to other dimensions? After all, physical attractiveness is somewhat defined by whether other people agree that you are desirable. It is the case that similar results are found for other dimensions of the self. For example, in another study, students on the first day of class were asked to rate their own personality and the personalities of the other students in the class. When students agreed that someone appeared friendly, that person tended to rate themselves as friendly (Norman & Goldberg, 1966). Finally, it is notable that there is evidence that similar neural systems support making self-judgments and imagining how one is perceived by other people (Ochsner et al., 2005; Pfeifer et al., 2009). Although future research is needed to understand whether similar covariance between neural regions predicts self-evaluation and imagining how one is perceived by other people, the commonality between these processes raises the possibility that they share at least some of the same psychological mechanisms. Therefore, existing research points to the strong likelihood of a close relationship between self-perceptions and how the self is perceived by other people.

A more recent incarnation of the looking-glass self perspective can be found in research on the sociometer hypothesis (Leary, Tambor, Terdal, & Downs, 1995). Specifically, this perspective states that self-worth fundamentally signals our acceptance by other people. In other words, self-esteem helps us bypass the process of imagining how other people see us and quickly signals the degree to which the self is positively (or negatively) viewed by others. In support of this theory, research has shown that self-esteem is determined both by how others actually feel about the self and perceptions of other people's regard.

For example, one longitudinal study examined the relation between self-perceptions, other people's perceptions of the self, and perceived social acceptance (Srivastava & Beer, 2005). Participants were randomly assigned to groups and performed various group tasks over the course of four meetings. Participants rated themselves, their group members, and how they believed they were perceived by their group members on various attributes. The more participants were rated as likable by their group members in the first group meeting, the more their self-perceptions of likability rose in subsequent meetings. However, this effect was independent of the effect of perceived regard on self-perceptions. In other words, self-perceptions of likability appear to increase when the self is well received by other people, but this effect is not wholly accounted for by knowledge of how the self is received by others. It is important to note that this study also examined the opposite possibility, that is, whether people broadcast aspects of themselves that then influence how they are perceived by other people. However, no significant effects were found for initial self-perceptions predicting other perceptions in later meetings. In contrast to the large body of correlational research demonstrating agreement between the self and others, this study reinforces the likelihood that those correlations may indeed reflect self-perceptions that are influenced by how we imagine other people to see us. In Jamesian terms, the looking-glass self perspective suggests that the perceiver self uses social reputation to learn about the perceived self.

Social Comparison Theory

The first two theories of self-knowledge acquisition have considered our ability to either draw on our own observations or other people's observations to learn about the self. A third perspective, social comparison theory, focuses on the role of other people as a reference point for evaluating the self (Festinger, 1954; Goethals & Darley, 1977; Kruglanski & Mayseless, 1990; Mussweiler & Rueter, 2003). From a Jamesian perspective, social comparison theory suggests that the perceiver self learns about the perceived self by comparing the perceived self to other people. For example, if you want to get a sense of your musical ability, you might compare your musical ability to the musical ability of your peers.

Times of uncertainty are particularly likely to bring out social comparisons. (See Kruglanski & Mayseless, 1990.) Furthermore, we make social-comparative judgments spontaneously and automatically (Gilbert, Geisler, & Morris, 1995) and, in fact, tend to identify people who are a part of our everyday lives who can serve as relatively chronic reference points (Mussweiler & Rueter, 2003). The power of social-comparative information is so great that it is difficult to discount it even when someone is aware that it is not relevant for self-evaluation. For example, participants under mental load could not help but reference the performance of another person even when it was clearly not relevant for self-evaluation (Gilbert et al., 1995). Participants viewed a videotape of a confederate performing a personality impression task. Participants were explicitly instructed that the confederate's task performance was due to external factors. Participants who viewed a confederate with high task performance were told that the confederate was performing the task for a second time. If the confederate had done poorly, they were told that the confederate had been given misleading information about how to perform the task. Afterward, each participant then performed the same personality impression task. Participants then rated their own competence at the task while maintaining an 8-digit number in their minds (or not). For participants who were not under mental load, their own competence was not affected by whether the confederate had done well or poorly on the task. However, participants in the mental load condition tended to rate themselves as if the confederate were an appropriate benchmark for average performance on the task. Even though they were told that the confederate had advantages or disadvantages they did not have, self-competence was lower when the confederate had done well rather than poorly. These results suggest that we automatically want to use salient social targets as referents for self-evaluation and that it takes extra cognitive resources to discount them.

Furthermore, we are loath to let our chronic points of reference go. Social comparisons often involve a referent that has been repeatedly useful in the past even when that person is not an optimal reference point for a particular comparison. For example, we may repeatedly compare ourselves to our friends such that we eventually tend to use them as a routine referent in our social-comparative judgments. In a series of studies, researchers found that self-evaluation and information about a friend tend to facilitate one another (Mussweiler & Rueter, 2003). Participants were faster at recognizing the name of their best friend after making a self-evaluation judgment of a personality trait (when compared to making a personality trait judgment for a celebrity). Furthermore, participants were also faster at judging the personality of their friend after making a self-evaluation judgment of personality trait (when compared to making a personality trait judgment for a celebrity). Importantly, these results held even when the self and best friend did not share the personality characteristic. Taken together these studies illustrate people's inclination to make social comparisons and that certain people become such routine referents that probing for self-evaluation increases accessibility of information about those other people.

Summary

People gain knowledge about themselves in myriad ways. They may observe their own behavior just as they might observe someone else's behavior as an avenue for making inferences about intangible, internal states. Additionally, other people play at least two roles in our self-evaluation processes. We may learn about ourselves or be particularly likely to internalize our behavior when we consider the perspective of an external observer. We also learn about ourselves by comparing ourselves to other people. In particular, social comparisons appear to be a particularly automatic way in which people gain knowledge about themselves.

MENTAL REPRESENTATIONS OF THE SELF

Once we have acquired information about the self, how do we represent it? Could we answer this question simply by reviewing the basic principles of knowledge representation? Or is there something special about the representation of self-knowledge because it is the only case in which we are both perceiver and perceived? (See Keenan & Baillet, 1980; Kihlstrom & Cantor, 1984.) Self-knowledge does not draw on dramatically different principles, yet it does highlight the effect of intimacy and frequency on knowledge representation. Mental representations of self-knowledge are particularly elaborate and well organized. The unique ways in which we create and access representations of self-information are illustrated by two effects: the self-reference effect and the relation of episodic memory to person judgment.

Self-Reference Effect

The self-reference effect is the tendency for people to remember information much better if it is encoded in relation to the self rather than to other referents. Are you athletic, dependable, and cheerful? Is the prime minister attractive, youthful, and liberal? Which of these words has more than two syllables: obnoxious, curious, furious? Now pretend you were in an experiment answering all of these questions, and 10 minutes later the experimenter gave you a surprise recall test. Your response would be characterized by the self-reference effect, that is, you would have better memory for the words you judged in relation to yourself than to other referents such as a politician or syllabic structure (e.g., Kelley et al., 2002; Markus, 1977; Ochsner et al., 2005; Rogers, Kuiper, & Kirker, 1977). The self-reference effect provided an opportunity to understand whether mental representations of self differed from other mental representations. Specifically, researchers wanted to know if the superior memory for self-reference information was the result of a unique cognitive process or an optimized case of the factors known to promote memory in other circumstances.

Schemas are theorized to aid people in organizing and guiding knowledge (e.g., Bartlett, 1932; Neisser, 1967; Taylor & Crocker, 1981), and self-knowledge has been conceptualized in terms of a self-schema (Markus, 1977). From this perspective, a self-schema would help organize the knowledge gained by observing the self, imagining other people's perspectives, and social comparison. Self-schemas tend to include information about the self that is important or central to defining the self (Markus, 1977). Schemas organize the information we currently hold; furthermore, they form the lens through which new schema-relevant information is processed (e.g., Baldwin, 1992;

Markus, Hamill, & Sentis, 1987; Taylor & Crocker, 1981). The self-reference effect is considered to reflect schematic influences on information processing.

The self-reference effect has been attributed to superior elaboration and organization of self-schemas (e.g., Ingram, Smith, & Brehm, 1983; Klein & Kihlstrom, 1986; Rogers et al., 1977; Symons & Johnson, 1997). Self-schemata affect how much information is elaborated during encoding, and information that is more extensively elaborated tends to be better remembered (Craik & Tulving, 1975). The superior memory for self-reference information has led researchers to theorize that the self-schema is especially well developed compared to other schemas (e.g., Markus, 1977; Rogers et al., 1977). Additionally, the rich nature of the self-schema should create many avenues to elaborate on the information being encoded in relation to the self. For example, if we ask you to judge whether you are talented, any number of self-associations may be triggered because there are so many domains in which someone could be talented. The encoding process becomes that much more elaborate because it is analyzed in relation to diverse preexisting self-information. However, if we ask you if the word "talented" has two syllables, you are not likely to process the question in terms of its meaning, let alone in a diverse manner. Instead, you will assess the syllabic content with a brief analysis of the word's pronunciation.

Just as self-schemas promote elaboration, they also are theorized to provide superior organization of information (Klein & Kihstrom, 1986; Symons & Johnson, 1997), and organization of information promotes memory (Bower, Clark, Lesgold, & Winzenz, 1969). For example, say we ask you to remember the list: dumbbell, mushroom, screwdriver, helmet, spinach, and saw. You will find it easier to remember the words if you organize them into three categories: sports equipment, vegetables, and tools (Bower et al., 1969). Now apply this logic to the self-reference effect. Researchers suggest that evaluating information in relation to the self organizes the information into categories (e.g., "me" and "not me"). For example, one study asked people to judge a list of body parts in relation to the self ("Can you think of an incident in which you had an injury or illness associated with your neck?") and to judge a different list of body parts on the extent to which the body parts were internal or external (Klein & Kihlstrom, 1986). In contrast to the typical superiority of memory for information in a self-reference condition, memory for the words was not significantly different across the two conditions. In other words, the memory advantage for words encoded in relation to the self is similar to words encoded in relation to parallel organizational cues (e.g., categories of me versus not me or internal versus external body parts). These studies suggest that we take extra pains when representing the information we gather about ourselves. Of all of the schematic representations we tend to create, the perceived self is represented with especially elaborate and well-organized schemas.

Role of Abstract and Episodic Information in Self-Judgment

Another aspect of mental representations of self that may be unique is the relation between abstract and episodic knowledge. Specifically, research suggests that we favor abstract information about the self when compared with other social targets (e.g., Klein, Babey, & Sherman, 1997; Klein, Loftus, & Burton, 1989). For example, if we ask you to decide whether the president is funny, you are likely to form your judgment by searching through your memory for instances that confirm or dispute the president's humor. But if we ask you to decide whether you are funny, your answer is not likely to involve a search through autobiographical memories.

Why would this process be so different for the self than for others? The sheer amount of time we spend with ourselves may lend itself to creating summaries or abstractions of episodic information about ourselves, but we may be less likely to do that for others or have fewer opportunities for creating abstract information about others (Klein, Loftus, Trafton, & Fuhrman, 1992). When judging ourselves, the availability of abstracted information may undermine the need or preference for autobiographical memories to draw conclusions about ourselves (e.g., Klein et al., 1989, 1997). In fact, research has shown that self-description judgments are not facilitated by recalling episodic information about the self. In one study, participants judged personality trait words for their self-descriptiveness (Klein et al., 1989). Before each judgment, participants performed one of three tasks: They generated a definition of the personality trait word, remembered a time they exhibited the personality trait, or made a self-descriptive judgment. If we make self-descriptive judgments by computing our answers from autobiographical memories, participants should have been faster when making a self-descriptive judgment after recalling an autobiographical memory than after generating a definition of the trait word. However, that is not what the study found. Instead, participants were equally quick to make self-descriptive judgments when generating a trait definition or when recalling an autobiographical memory. Furthermore, when participants performed one of these tasks before being asked to recall an autobiographical memory, they were no quicker to do so after recalling an autobiographical memory than after generating a semantic definition (Klein et al., 1989). These studies show that the processes involved in self-description judgments and retrieving autobiographical memories are not redundant.

However, the study of self-descriptive judgments in new contexts shows a different pattern. Specifically, personality judgments of the self in new contexts do rely on autobiographical memories. For example, participants were asked to perform the tasks just mentioned but either in relation to a context in which they had long-term experiences or a context in which they had short-term experiences (Klein et al., 1992). Specifically, participants were asked to recall memories and make self-description judgments in relation to the way they acted at home with their families (i.e., long-term experience) or in relation to the way they acted in college (i.e., short-term experience). All participants had only about 3 months of experience with their college environments (compared to 18 years of experience in the family environment). The results for judgments made in relation to home matched the results described earlier: Autobiographical recall did not significantly facilitate self-description judgment. However, recalling autobiographical instances from time spent at college did facilitate self-description judgments in the college context. In other words, participants were faster at making decisions about whether a trait described them if they retrieved a specific trait-relevant memory when evaluating their self in a new context.

In some sense, these findings are consistent with self-perception theory, which emphasizes that we learn about ourselves by observing our behavior. The research on the facilitative effect of autobiographical memory on self-description suggests that this theory is particularly relevant for forming impressions of the self in new contexts. Finally, making personality judgments without relying on episodic memory may be mostly unique to the self. For the most part, retrieving episodic memories about one's mother will facilitate personality judgments of her (Klein et al., 1992).

Summary

Taken together, the research on the self-reference effect and the role of memory retrieval in self-judgment illustrates the unique ways in which we mentally represent self-knowledge. In comparison to other kinds of knowledge, we represent knowledge about ourselves in particularly elaborate and well-organized ways. The rich representation of self-knowledge also has implications for how we make self-description judgments. In contrast to other social targets, even close social targets, such as a parent, our extensive experience with ourselves may create abstract representations culled from repeated experiences. As a result, self-description judgments often do not necessitate the retrieval of autobiographical memories unless they are from a relatively new context.

MOTIVATION AND SELF-KNOWLEDGE

So far the theories of self-knowledge acquisition have described different avenues for learning about the self, but they have not spoken to whether these processes operate in an amotivational or a motivated manner. After all, we do have some choice over the behaviors we focus on, whose opinion we internalize, and who we view as a relevant referent for comparison. These choices will have implications for whether we are likely to see ourselves in positive or negative light. Do we act to ensure a particular view of ourselves? Yes, we often do. In fact, a central theme in diverse disciplines such as psychology, philosophy, and economics is that we are rarely dispassionate when it comes to self-processes. Therefore, the next set of questions we need to address is about the motivations that may affect our self-knowledge acquisition. What kinds of motivations influence self-evaluation processes and the application of self-knowledge? Does a particular motivation dominate all others? Are motivational influences always at work, or are there any circumstances in which we tend to remain dispassionate when evaluating ourselves? Research consistently has found evidence for at least three broad motivations that influence self-perception processes, and many factors affect which motivation is likely to prevail over the others (e.g., Anderson, Srivastava, Beer, Spataro, & Chatman, 2006; Kwang & Swann, 2010; Sedikides & Gregg, 2008; Swann et al., 1990; Taylor & Brown, 1988; Trope, 1980): self-enhancement, self-verification, and self-assessment.

Self-Enhancement

Self-enhancement is the motivation to see oneself in a positive light. It is considered to be so ubiquitous and fundamental by some researchers that they have equated it with the need for food (e.g., Alicke, Klotz, Breitenbecher, Yurak, & Vrendenburg, 1995; Paulhus, Graf, & van Selst, 1989; Sedikides & Gregg, 2008; Taylor & Brown, 1988). A large body of research has shown that people often evaluate themselves as though they are wearing rose-colored glasses, that is, in a positively skewed manner. People claim to have more positive traits than negative traits, believe they have done better on a task than indicated by their actual performance, rate themselves as having more desirable personalities than their peers, and attribute their failures to circumstance rather than essential qualities of the self (e.g., Alicke et al., 1995; Beer & Hughes, 2010; Jones & Nisbett, 1971;

Klayman et al., 1999; Robins & Beer, 2001; Sedikides & Gregg; 2008; Taylor & Brown, 1988; Taylor & Koivumaki, 1976).

When unrealistically positive self-perceptions are considered from a motivational perspective, researchers theorize that they serve to maintain or inflate self-worth (Sedikides & Gregg, 2008; Taylor & Brown, 1988; but see Chambers & Windschitl, 2004). We feel good about ourselves when we see ourselves in a positive light. We can accomplish this by influencing what we observe about ourselves, how we compute social comparisons, and imagine how others see us. For example, we can store positive information about ourselves in a more accessible manner than negative information, choose social comparison referents who guarantee a favorable comparison, and select interaction partners who provide positive feedback (Swann, Hixon, Stein-Seroussi, & Gilbert, 1990).

For example, research has shown that people's observations of themselves are represented in such a way that positive information is more accessible than negative information. Studies of mental load have shown that people automatically report having significantly more positive qualities and fewer negative qualities than other people (e.g., Beer, Chester, & Hughes, 2013; Beer & Hughes, 2010; Paulhus et al., 1989). People are also unlikely to remember negative feedback about themselves as well as positive feedback, and this poor memory may result from negative information being stored separately from one's self-concept (Pinter, Green, Sedikides, & Gregg, 2011). These studies suggest that positive information in self-schemas is more easily accessible than negative information.

Additionally, people choose referents for social comparison that guarantee that they always will compare favorably. For example, people may idiosyncratically define a personality trait so that they seem exceptional when compared to others (Alicke et al., 1995; Dunning, Meyerowitz, & Holzberg, 1989). If you want to see yourself as a good cook compared to your friends, you can define a good cook as someone who excels at things you can do (e.g., never burn the food) but downplay the importance of things that you do not do (e.g., developing unique recipes). Additionally, people can enhance their self-worth by comparing themselves to people who are worse off (e.g., Kruglanksi & Mayseless, 1990). For example, students in a medical training program ensured positive self-appraisals of their performance by choosing to compare themselves to peers with poorer performance (Buunk, Cohen-Schotanus, & van Nek, 2007).

In summary, self-enhancement motivation may create lopsided representation or greater elaboration of positive information about the self (compared to negative information about the self). We may evaluate ourselves by observing our behavior or imagining what others think of us, but we do not do this dispassionately. Instead, there is evidence that we ensure positive self-views by focusing on the positive aspects of our behavior and the good things that people have to say about us. Additionally, we ensure positive self-views by comparing ourselves on dimensions or to people that emphasize our positive qualities.

Self-Verification

Another motivation that influences the acquisition and representation of self-knowledge is self-verification (e.g., Kwang & Swann, 2010; Swann et al., 1989). The central tenet of self-verification theory is that people strive to confirm their current views of themselves. The motivation to feel that the self is consistent is one means of comprehending the world at large. To the extent that self-views are consistent and predictable, people may feel that the world is predictable and coherent, which, in turn, gives them a sense of control.

How can self-verification and self-enhancement be distinguished from one another? The majority of people have moderate to high levels of self-esteem (e.g., Gray-Little, Williams, & Hancock, 1997). Therefore, verification of these views would manifest in the positively tinged evaluations typically associated with self-enhancement motivations. The consideration of people with low self-esteem highlights the key difference between the two theories. Self-verification would be indicated by situations where people with low self-esteem strive to confirm their negative self-views. Decades of research have confirmed this hypothesis: People with low self-esteem perpetuate their negative self-views by seeking information and environments that reinforce their negative self-views (see Kwang & Swann, 2010). The finding that people find it easier to remember positive versus negative information about the self (Pinter et al., 2011) is qualified by research showing that people find it easier to remember information based on its consistency with their current self-view. For example, participants who varied in self-esteem were given false feedback about the desirability of their personality (Story, 1998). Participants were more successful at recalling the feedback if it matched their own self-views. The critical support for self-verification theory was the finding that participants with low self-esteem found it easier to remember details of their feedback if it indicated they had undesirable personalities.

These findings cannot be explained by something that is categorically different about people with low self-esteem. These same effects are evident when people with high self-esteem face feedback about their perceived flaws (Swann et al., 1989). Participants were prescreened to ensure that while they had positive self-views about some aspects of their personality, they also held negative self-views about other aspects of their personality. Participants were then led to believe that they had been evaluated by three people (who were confederates) and asked to rate their preference for interacting further with each person. The evaluation feedback presented to participants was bogus; it was manipulated to verify either positive or negative aspects of self-ratings of personality. All participants, regardless of high or low self-esteem, preferred to have future interactions with people they believed to share their perceptions of both their positive and their negative attributes. This research illustrates the way in which people seek and recall information that reinforces current self-views even when they are negative. Self-verification operates on our acquisition and representation of self-knowledge when we focus our self-observations on information that reinforces our current self-views and when we choose to interact with people who will reflect those views back to us.

Self-Assessment

A third motivation has been proposed that captures the dispassionate search for self-information in the interest of self-assessment. This theory finds support in research showing that there are times that people do gather information about themselves in order to gain accurate self-knowledge (e.g., Trope, 1986). People sometimes are willing to forgo the good feelings of self-enhancement and control provided by self-verification to gain an accurate understanding of their capabilities and how to improve on them. (Some researchers treat this latter benefit separately from self-assessment by delineating it as self-improvement: e.g., Taylor, Neter, & Wayment, 1995.) People prefer to receive feedback compared to no feedback about themselves (e.g., Dunning, 1995), and interest in that feedback increases to the extent it is believed to hold objective, diagnostic value (Trope, 1975). Interest in this kind of feedback persists even when we know it may undermine self-esteem (Trope, 1980).

For example, in one study, baseline measurements of achievement motivations were collected in order to understand how individual differences related to preference for accurate feedback (Trope, 1975). Several weeks later, participants completed a series of tests that varied in difficulty and learned about the extent to which each test had the potential to accurately assess their abilities. Participants, especially those with high achievement motivation, were most interested in feedback from tests that were presented as highly diagnostic. Furthermore, participants did not simply want to know how they had performed on the easy tests. Research on self-assessment demonstrates that people are interested in gathering feedback even when it may not enhance or verify current self-views. From this perspective, a relatively nondefensive curiosity about the self governs at least some of the data gathering and mental representation associated with self-evaluation.

Relative Influence of Each Motivation

Research has shown that a number of motivations can operate on self-evaluation processes. What do we know about when each motivation is likely to operate? Is one motivation more prevalent than the others, or is the influence of each of the motivations determined by domain or some other factor? Some researchers have suggested that self-enhancement may be more automatic and affect-driven than self-verification and self-assessment whereas self-verification and self-assessment may require more controlled and cognitive processing (e.g., Swann et al., 1990; Trope & Neter, 1994).

As mentioned, people's automatic tendency often is to seek self-enhancing information. When people's cognitive resources are undermined, such as when they are under cognitive load, their self-description judgments and social comparisons become even more positively skewed (e.g., Beer & Hughes, 2010; Beer et al., 2013; Paulhus et al., 1989). When under mental load, people are more likely to select social interaction partners who enhance their self-view rather than verify their self-view (Swann et al., 1990).

However, the automatic tendency toward self-enhancement should not be taken as evidence that it somehow always dominates self-evaluation. In fact, there are even certain domains in which self-evaluations typically are not enhanced. For example, on average, people do not inflate their self-perceptions of status. A series of studies examined self-perception of status in experimentally assigned groups and in naturalistic groups (Anderson et al., 2006). People were aware of their status; their self-ratings correlated significantly with group member ratings. Furthermore, the correlation between self and group member ratings of status holds for minimally acquainted groups and across time. In contrast to studies in other domains where self-perceptions tend to be more positive than objective indicators, status is one domain in which self-enhancement is not expressed on average.

An open question is whether domains in which self-enhancement is not expressed reflect a failure to automatically engage that motivation in favor of others or whether it is attenuated through cognitive effort. This possibility is raised by research showing that in other domains, people use cognitive effort to accomplish self-verification and self-assessment over self-enhancement. When not undermined by mental load, people seek interaction partners who verify their negative self-view (Swann et al., 1990). And cost-benefit analyses tend to precede the commitment to self-assessment; people are most likely to engage in self-assessment when they know that learning something negative about the self ultimately will be useful despite adverse effects on self-esteem (Trope & Neter, 1994).

Factors other than automaticity have been investigated for their role in determining the relative influence of the various self-perception motivations. One possibility is that the congruence of a domain with the goal of a motivation influences its effect. This possibility has been investigated, but there is not strong evidence for it. For example, if people are motivated to assess themselves accurately in order to promote achievement, then we might hypothesize that self-assessment will show a dominant influence on perceptions of achievement-related qualities, such as academic ability. However, research has found that people, on average, tend to self-enhance their academic ability (e.g., Robins & Beer, 2001). Furthermore, some research suggests that we are just as likely to respond to the bruises of past and future self-esteem threats with not just self-enhancement but also self-verification or self-assessment. For example, in one study, participants described the situations in which they had been motivated to enhance, verify, assess, or improve themselves through their self-evaluations (Taylor et al., 1995). The descriptions were submitted to a narrative analysis which found that situations of threat, in either the past or the future, were equally likely to elicit the different motivations. However, more recent research has found that people self-enhance less when recalling a past self-esteem threat when compared to current self-esteem threats (Gramzow & Willard, 2006). Therefore, it is possible that congruence between motivation and domain can play a role in determining when motivations are likely to operate, but the effects do not appear to be robust across studies.

A second possibility is raised by James's focus on the social nature of the self. Perhaps certain relationships tend to elicit particular motivations for self-evaluation. Consistent with this perspective, a recent meta-analytic review suggested that self-verification motivations may be particularly strong when seeking information in long-term relationships compared to newly forming relationships (Kwang & Swann, 2010). Taken together, all of these findings point to the need for more research aimed at understanding how and when self-evaluation processes and representation of self-knowledge are influenced by various motivations.

NEURAL REPRESENTATIONS OF THE SELF

Research on the self has recently made important strides toward understanding the neurobiological basis of self-processes. An interest in self-processes took root in neuroscience circles when neuropsychologists repeatedly noted that frontal lobe damage often was associated with impaired self-insight (Blumer & Benson, 1975). However, the emergence of neuroimaging techniques has led to a body of empirical data to complement the observation data. In particular, recent lesion studies and neuroimaging in healthy populations suggest that different aspects of self-evaluation draw on different frontal lobe subregions. More specifically, this work suggests that the neural systems which mediate the mental representation of self-knowledge are distinct from the operation of self-perception motivations.

Self-Reference Encoding

The first major wave of empirical investigations of the neural basis of self-process focused on the self-reference effect. A large body of literature has found robust, convergent evidence that the medial prefrontal cortex (mPFC) plays a role in encoding and remembering self-reference information. Neuroimaging studies have drawn on positron emission tomography and functional magnetic resonance imaging methodologies to measure neural activation while participants perform the self-reference paradigm typically used in behavioral research. For example, participants might evaluate personality words for their self-descriptiveness, descriptiveness of a political figure, general social desirability, and number of syllables. These experiments find that evaluating personality trait words in relation to the self (compared to the conditions just mentioned) tends to increase activation in the mPFC (Brodmann's area 9/10) (Craik et al., 1999; Fossati et al., 2003; Gillihan & Farah, 2005; Kelly et al., 2002; Kircher et al., 2002; Ochsner et al., 2005). In fact, the mPFC activation increases as a function of self-descriptiveness; the magnitude of its activation correlates with the extent to which the personality trait is considered to be self-descriptive (e.g., Kircher et al., 2002; Macrae, Moran, Heatherton, Banfield, & Kelley, 2004; Moran et al., 2006).

Furthermore, the association between self-reference and mPFC activation is not limited to personality judgments. Research has found a similar association between mPFC activation and self-reference faces (Keenan, Wheeler, Gallup, & Pasucal-Leone, 2000; Kircher et al., 2000) and objects (Kim & Johnson, 2012). For example, participants might observe their own face and the faces of a novel person or a close other. Observations of one's own face (in comparison to the face of another person) are associated with increased activation in the right frontal lobe (e.g., Brodmann's area 9/10) (e.g., Keenan et al., 2000; Kircher et al., 2000). Another study asked participants to encode objects either as owned by the self (self-reference) or by someone else (Kim & Johnson, 2012). Specifically, participants were presented with objects and given rules about when they should place the objects in a basket assigned to them or another person. Participants also were asked to imagine that they owned the objects that they placed in their baskets. Objects assigned to the self condition were associated with increased mPFC activation (compared to objects assigned to the other person).

Although the majority of neural studies on the self-reference effect have not taken the step of understanding the relation between the mPFC activation and the superior memory effect, a few studies do find this relation (e.g., Kim & Johnson, 2012; Macrae et al., 2004). For example, one study asked participants to rate personality trait words for their self-descriptiveness and then gave participants a surprise memory test (Macrae et al., 2004). Activation in the mPFC predicted which words were later remembered compared to those words that were not remembered. Additionally, the relation between mPFC and superior memory for self-reference information extends to objects. In the transient ownership study mentioned earlier, participants were given a surprise memory test where they had to remember which objects they had placed in their own basket versus the basket of another person. Medial prefrontal cortex activation predicted which of the objects assigned to the participant's basket would be later remembered (Kim & Johnson, 2012).

Motivated Self-Perception

In contrast to the large body of literature on the neurobiology of self-reference encoding, less attention has been paid to the neurobiology of motivated self-perception (Beer, 2007, 2014). Only self-enhancement has received any sort of consistent attention in the neural literature. Current research shows that there are two neural profiles of self-enhancement (Beer, 2014; Flagan & Beer, 2013). One set of lesion and neuroimaging studies has found that self-enhancement is associated with reduced orbitofrontal cortex (OFC) activation; these studies have operationalized self-enhancement in a number of ways including self-evaluations that diverge from objective indicators (e.g., actual performance, base rates). However, these studies have not manipulated the presence of self-esteem threat. In the studies that do manipulate self-esteem threat as a means of eliciting self-enhanced responses, increased OFC activation is associated with self-protective self-evaluations (Hughes & Beer, 2012a, 2013). Furthermore, the OFC is operating within separate neural networks in the studies in which self-esteem threat is manipulated (Flagan & Beer, 2013).

The earliest investigations of the neural basis of self-enhancement were rather indirect. Specifically, researchers were focused on identifying neural regions that tracked whether self-judgments involved socially desirable information or not. For example, participants were asked to rate the self-descriptiveness of desirable and undesirable personality traits (Beer & Hughes, 2010; Moran et al., 2006) or to evaluate the likelihood that good and bad events would happen to them in the future (Sharot et al., 2007). Studies taking this approach consistently found that the ventral anterior cingulate cortex (vACC) differentiates desirable attributes from undesirable attributes. Specifically, vACC activation increases when people rate desirable compared to undesirable personality traits and when they evaluate the likelihood that they will experience good events in the future compared to bad events. However, claiming that a desirable attribute is self-descriptive does not necessarily indicate the influence of an active self-enhancement motivation (see Beer, 2007; Beer & Hughes, 2010; Chambers & Windschitl, 2004). Measuring self-enhancement with claims of positive attributes is problematic because (a) people may genuinely be characterized by a desirable quality and (b) self-enhancement has been shown to involve both the inflation of desirable attributes and the dismissal of undesirable attributes (Beer & Hughes, 2010; Dunning et al., 1989; Taylor & Brown, 1988). Therefore, research involving more direct measures of the influence of self-enhancement motivation was needed to investigate the role of vACC.

The possible role of vACC was not supported in subsequent research. Instead, an emerging body of research has now shown that unrealistically positive self-evaluations tend to be associated with reduced OFC function rather than changes in vACC activation. The relation to reduced OFC function holds when unrealistically positive self-evaluations are operationalized as discrepancies between self-confidence and actual task performance (Beer, Lombardo, & Bhanji, 2010), base rates compared to self-rankings in social comparisons (Beer & Hughes, 2010), attributions for task success compared to task failure (Blackwood, Bentall, Simmons, Murray, & Howard, 2003), and self-compared to other perceptions (Beer et al., 2006).

For example, self-enhancement is implicated when self-evaluations diverge from objective indicators, such as task performance. Overestimation of success on a trivia task is associated with reduced OFC activation (Beer et al., 2010). Participants answered trivia questions about average July temperatures in U.S. cities and then estimated their confidence that their answers were correct. When participants had answered the actual trivia question incorrectly, a region of medial OFC was negatively modulated by confidence level. In other words, for those incorrect trials where confidence was unwarranted, people tended to recruit OFC activation less often. It is important to note that the relation could not be explained by confidence level alone; OFC did not predict confidence for trials that were answered correctly. Additionally, participants who tended to be more overconfident about their performance on the task were the least likely to activate OFC.

Another indicator of self-enhancement is when people compare themselves in an unrealistically positive manner to their peer group. People tend to believe they have significantly more desirable personality traits and significantly fewer undesirable personality traits than their peers (e.g., Dunning et al., 1989). Although each person is likely to have a unique comparative ranking on some traits, so is the average peer. Therefore, ranking the self as having significantly more desirable traits and fewer negative traits is considered a marker of self-enhancement (Taylor & Brown, 1988; but see Chambers & Windschitl, 2004). OFC activation is reduced when people make unrealistically positive social comparisons compared to social comparisons that are more realistically calibrated (Beer & Hughes, 2010). Participants were asked to compare themselves to their average peer on 200 personality traits (100 desirable traits, 100 undesirable traits). The more participants rated themselves as having both more desirable traits and fewer negative traits than their average peer, the less likely they were to activate OFC during the social-comparative judgments (Beer & Hughes, 2010).

Another classic example of self-enhancement is when people make unrealistically positive attributions for their behavior. Specifically, they tend to take credit for their successes but disavow responsibility for failure (Taylor & Brown, 1988). Reduced OFC activation is associated with people's choices to account for their behavior in this self-serving manner (Blackwood et al., 2003). In one study, participants imagined that they had experienced social success (i.e., a friend gives you a gift) or social failure (i.e., a friend refuses to talk to you). Participants then had to report on whether they had imagined the success or failure in such a way that they were responsible or someone else was responsible. The trials in which participants attributed their imaginary success or failure to self-serving factors (i.e., self for success, friend or situation for failure) were compared to non-self-serving factors (i.e., self for failure, friend or situation for success). Taking credit for success and dismissing self-responsibility for failure was associated with reduced OFC activation.

Finally, findings from lesion research are consistent with the neuroimaging research. Specifically, unrealistically positive evaluations of one's task performance (in comparison to judges' ratings) are associated with OFC damage. For example, one study found that patients with OFC damage overestimated their social skills on a social interaction task when compared to patients with lateral PFC damage or healthy control participants (Beer et al., 2006). Participants had to engage in a semistructured conversation with the experimenter who was a stranger to them. All participants self-reported that they were aware that social norms dictate that conversations with strangers follow certain rules of discretion. However, patients with OFC damage were likely to introduce personal information into the conversation that was more suitable for a conversation with an intimate other. Self-reports of social appropriateness were collected after the conversation ended. Compared to the healthy control participants and other patients with lesions, patients with OFC damage had poor insight into the inappropriateness of their conversation (when compared to perceptions of blind judges).

However, the neural profile of self-enhanced responses is different when the self-evaluations are a response to an immediate threat to self-esteem. For example, when the social-comparative paradigm used in an earlier functional magnetic resonance imaging study (Beer & Hughes, 2010) was modified and combined with a self-esteem threat manipulation, a different neural profile predicted downplayed rankings of negative traits. Specifically, participants made social-comparative judgments after they had received feedback that their peers either did not find them attractive (i.e., a threat condition) or did find them attractive (Hughes & Beer, 2013). When participants received threatening social feedback, they were significantly more likely to downplay their negative traits in comparison to an average peer. Individual differences in self-flattering social comparisons were positively associated with OFC activation (Hughes & Beer, 2013). A follow-up study also found that self-protective self-evaluations were associated with increased OFC activation (Hughes & Beer, 2012a). Specifically, OFC activation predicted the extent to which participants ensured that they would not be exposed as frauds when performing an intelligence test. A secondary analysis of these data sets revealed that the region of increased OFC activation associated with self-enhancement tends to covary with middle frontal gyrus and caudate whereas the OFC region of decreased activation associated with self-enhancement operates in a widespread network (Flagan & Beer, 2013). Taken together, this research finds consistent evidence that there are two different neural profiles of self-enhancement: one involving a network centered on decreased OFC activation and another involving a network centered on increased OFC activation.

BRIDGING THE NEUROSCIENCE AND PSYCHOLOGICAL RESEARCH ON THE SELF

The neural research on self-processes has made important advances for neuroscience by delineating more specific relations between frontal lobe function and self-processes, but does it have any implications for deepening our understanding of the psychology of these processes? Future research that directly investigates the psychological significance of the neurobiological findings can help develop psychological theory in addition to neural models. However, two intriguing possibilities are already emerging from the current research.

First, the research associating vACC activation with rating socially desirable attributes raises the possibility that we are spontaneously vigilant for opportunities to cast ourselves in a positive light. The existing research shows that vACC activation does not predict self-enhanced responses. However, it may be sensitive to opportunities where self-enhancement has the potential to be fulfilled. In the studies reviewed earlier in the chapter, vACC differentiated conditions of desirable traits from undesirable traits, even though participants were not asked to evaluate the traits for their desirability, just their self-descriptiveness. Subsequent research also has shown that the extent to which vACC differentiates desirable conditions from undesirable conditions depends on how much we care about viewing the target of the judgment in a positive light. That is, vACC is especially likely to differentiate desirable from undesirable attributes when we are judging attributes for people we want to cast in a positive light, such as ourselves (Moran et al., 2006), and other people we care about, such as romantic partners (Hughes & Beer, 2012b). In other domains, vACC activation has been associated with detecting the potential for reward (e.g., Rogers et al., 2004). Therefore, it is possible that the pattern of vACC activation seen in these studies of social cognition reflects its role in detecting which judgment conditions are likely to be rewarding (i.e., flattering), and this function is especially engaged when judging the self or people we want to cast in a positive light. If this were the case, it suggests that one way the motivation to self-enhance influences information processing is through vigilance for opportunities where self-enhancement could be fulfilled.

Second, the distinct neural profiles of self-enhancement suggest a possible reconciliation of the long-standing debate about the mechanisms underlying self-enhancement (Beer, 2014; Chambers & Windschitl, 2004; Taylor & Brown, 1988). The existing research finds fairly consistent evidence that different patterns of neural activation are associated with self-enhancement, depending on whether it is used as a response for coping with an immediate threat. Therefore, this neural data may explain why some researchers have described self-enhancement as a reflection of the desire to protect self-esteem (Taylor & Brown, 1988) whereas others have characterized it as the result of incomplete information processing (Chambers & Windschitl, 2004). The differential neural patterns of activation may indicate that different underlying mechanisms are engaged when self-enhancement arises from each of these causes even though they may appear similar at the level of a behavioral response (i.e., a skewed self-evaluation). In this case, both perspectives are correct. The possible psychological explanations of the research on the vACC and the distinct neural profiles of self-enhancement are just two examples of the psychological advances that can be achieved through neural investigations of self-processes. Future research is needed to build on these findings and raise new insights.

THE SELF OR PEOPLE IN GENERAL?

As this review has shown, a prevalent question in self research has been how much self-processing really represents unique operations or whether these processes extend to understanding ourselves as well as other people. In other words, researchers continue to wonder how special the self is. Surprisingly, researchers who study the self often conduct their research in parallel with researchers interested in social cognition about other people. Yet the literature suggests that there are many parallels in the underlying social cognitive processes. We also apply many of the expectations and heuristics that we use to evaluate ourselves to understanding other people (e.g., Knee et al., 2003; Malle, 2006; Nisbett & Wilson, 1977; Plaks et al., 2009). And as mentioned earlier, meta-analyses have shown that classically believed differences between self and other evaluation are not as robust or extensive as previously thought (Malle, 2006). The parallels in the underlying processes between self-perception and other perception is especially strong when the other person is someone we know intimately or someone we perceive to be similar to the self (e.g., Klein et al., 1992; Mitchell, Banaji, & Macrae, 2005; Mitchell, Macrae, & Banaji, 2006; S. L. Murray, Holmes, & Griffin, 1996; Neff & Karney, 2005; Symons & Johnson, 1997).

Highly Elaborated and Well-Organized Schema for Self and Close Others

For example, the rich elaboration and organization that is theorized to characterize self schemas likely is also present in the schemas for people close to us. Our schemas of close others may have rich elaboration and organization because we also enjoy a memory advantage for information encoded in relation to close others, although it is not quite as strong for self-reference encoding (e.g., Maki & McCaul, 1985; Ochsner et al., 2005; see Symons & Johnson, 1997, for a meta-analysis). In other words, if we ask you to judge whether a series of personality traits describes someone close to you (i.e., spouse, friend, child, sibling, or parent), then you are likely to remember these personality trait words almost as well as you would remember traits you rated in relation to yourself. And this memory would be even greater than if you judged the relevance of the trait words to a familiar, but not intimate, other person, such as a politician (e.g., Maki & McCaul, 1985).

Judgments About Both Self and Close Others Driven by Abstract Representation

The parallels in the way we make self-judgments and judgments about the highly descriptive personality traits of close others are also indicative of the development of rich schemas for close others. Just as we do not rely on the retrieval of episodic information to judge whether a personality trait describes us, we do not need to retrieve episodic information when we judge the personality traits we deem to be highly descriptive of people we know well. As mentioned previously, autobiographical memory retrieval tends to facilitate personality judgments of one's mother but not of the self. But there is one exception to this finding. The retrieval of memories does not facilitate judgments of the personality traits that are most characteristic of one's mother (Klein et al., 1992). These results suggest that we form abstract representations of the most central aspects of our close others' personalities.

Neural Similarities Underlie Representations of Self and Close Others

Finally, parallels between self-evaluation and other-evaluation exist at the neural level of analyses. There is at least some overlap in the neural systems that support self-evaluation and evaluation of close others (for reviews see Gilihan & Farah, 2005; R. J. Murray, Schaer, & Debbane, 2012; Ochsner et al., 2005) and similar others (Mitchell et al., 2005, 2006). Whereas significantly higher activation in mPFC is seen when self-evaluations are compared to evaluations of a familiar person, such as a politician (Kelley et al., 2002), this difference disappears when self-evaluations are compared to evaluations of close others, such as a romantic partner (e.g., Ochsner et al., 2005). Additionally, overlapping neural systems are found when self-evaluations are compared to evaluations of a relatively unknown person to the extent that the person is perceived as similar to the self (Mitchell et al., 2005, 2006). For example, one study required participants to evaluate the mental states or physical attributes of unfamiliar social targets. Participants viewed a series of faces and judged whether the face looked pleased (mental state) or symmetric (physical attribute). Afterward, participants rated the extent to which the people in the photographs were likely to be similar or dissimilar to the participants. Medial prefrontal cortex activation was associated with judgments about these strangers' mental states, and this activation increased to the extent that participants believed the strangers to be similar to themselves (Mitchell et al., 2005). The researchers suggested that this pattern of activation is consistent with the theory that, in the absence of other information, we may engage the self-system to evaluate novel others but only to the extent that the self seems like a reasonable means of understanding a novel person. Together these studies illustrate the neural and psychological commonality between self-evaluations, evaluations of other people who are intimate with the self, and evaluations of unfamiliar people who appear similar to the self.

Motivation Influences Perceptions of Both the Self and Close Others

The parallels between evaluation of the self and of close others also extend to the types of motivations that influence mental representations and evaluations. The desire to see the self in a positive light extends to close others (e.g., S. L. Murray & Holmes, 1997; Neff & Karney, 2005; Suls, Lemos, & Stewart, 2002; Taylor & Brown, 1988; Taylor & Koivumaki, 1976). The positive skew in evaluation is specific to close others; we tend to have more even-handed perceptions of the desirability of an unknown, typical person's personality (Suls et al., 2002). Similarly, we are much more likely to excuse away the poor social behavior of ourselves, our friends, and our spouses with attributions to situational factors but not the poor social behavior of strangers or people we dislike. For example, a series of studies asked participants to consider poor social behavior, such as showing up late to an appointment, having an argument, and ignoring others at a party (Taylor & Koivumaki, 1976). Participants rated the likelihood that dispositional and situational factors would motivate themselves, their spouses, and acquaintances to act in these ways. Situational factors were selected as explanations more often for the poor social behavior of the self and spouses.

Furthermore, similar neural regions support the flattering views of ourselves and the people we care about. Self-enhancement and enhanced evaluations of close others both draw on reduced OFC activation (Beer & Hughes, 2010; Hughes & Beer, 2012b). For example, in one study, participants compared their romantic partners and assigned roommates to an average peer on a series of 200 personality traits (100 desirable, 100 undesirable). OFC activation was reduced to the extent to which participants considered their romantic partners or roommates to have significantly more desirable personality traits and significantly fewer undesirable personality traits than the average peer.

Finally, self-evaluation and other-evaluation are influenced by more than just enhancement motivations. Just as we strive to confirm our self-view (e.g., choosing to interact with people who confirm our self view: Swann et al. 1989), we also structure environments to confirm our views of other people (e.g., Darley & Gross, 1983; Word, Zanna, & Cooper, 1974) even when we have not consciously accessed those views (Chen & Bargh, 1997). For example, people may impact a social interaction such that disliked others are set up to exhibit behavior that confirms negative evaluations. In one study, participants interviewed out-group members (compared to in-group members) for an ostensible job. They tended to cut the interview short, committed more speech errors, and were less invested in the interaction (Word et al., 1974). A follow-up study showed that these behaviors evoked poorer behavior from the interviewees. Subsequent participants were trained to use the poorer interview style to evaluate in-group applicants. Under those conditions, the in-group applicants came across as less competent and more nervous. These studies illustrate the ways in which motivational influences known to operate in self-evaluation can extend to evaluations of other people.

CONCLUSION

From a social cognitive standpoint, William James (1890) long ago pointed out that the self is an interesting case that involves both the perceiver and perceived. In the Jamesian conceptualization of the self, we are as much our physical presence as we are our innermost strivings and social reputations. We have discovered that we gather information about ourselves by observing the self through our own eyes and the eyes of others as well as comparing ourselves to others. Our self-knowledge tends to be represented in particularly elaborate and well-organized self-schemas. Although we have the capability and infrequent desire to gather and represent self-information in an even-handed manner, we are most often motivated to enhance or verify our self-views. Furthermore, there has been a recent movement to bridge the psychological and neural understanding of the self. We are just beginning to learn about how frontal lobe subregions contribute to different self-processes and the motivations that influence them. Finally, there is not strong evidence that self-evaluation is terribly unique. There are many parallels between the processes and motivations that shape self-evaluation and those that shape evaluations of other people, particularly those close to us or those we perceive to be similar to the self. Despite all of these advances, there is much left to learn about the self. For example, we have yet to understand how people balance competing motivational pressures on self-evaluation and the implications of the neural architecture of self-processing for understanding the underlying psychology.

REFERENCES

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, *55*, 387–395.

Alicke, M.D., Klotz, M.L., Breitenbecher, D. L., Yurak, T. J., & Vrendenburg, D. S. (1995). Personal contact, individuation, and the above-average effect. *Journal of Personality and Social Psychology*, *68*, 804–825.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychological Bulletin*, *111*, 431–441.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268–277.

Andersen, S. M., & Ross, L. (1984). Self-knowledge and social inference: I. The impact of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology*, *46*, 280–293.

Anderson, C. P., Srivastava, S., Beer, J. S., Spataro, S. E., & Chatman, J. A. (2006). Knowing your place: self-perceptions of status in social groups. *Journal of Personality and Social Psychology*, *91*, 1094–1110.

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin*, *112*, 461–484.

Bartels, D. M., & Rips, L. J. (2010). Psychological connectedness and intertemporal choice. *Journal of Experimental Psychology: General*, *139*(1), 49–69. doi:10.1037/a0018062

Bartlett, F. A. (1932). *A Study in Experimental and Social Psychology*. New York, NY: Cambridge University Press.

Baumeister, R. F. (1987). How the self became a problem: a psychological review of historical research. *Journal of Personality and Social Psychology*, *52*, 163–176.

Beer, J. S. (2007). The default self: feeling good or being right? *Trends in Cognitive Sciences*, *11*, 187–189.

Beer, J. S. (2014). Exaggerated positivity in self-evaluation: a social neuroscience approach to reconciling the role of self-esteem protection and cognitive bias. *Social and Personality Psychology Compass*, *8*, 583–594.

Beer, J. S., Chester, D., & Hughes, B. L. (2013). Social threat and cognitive load magnify self-enhancement and attenuate self-deprecation. *Journal of Experimental Social Psychology*, *49*, 706–711.

Beer, J. S., Heerey, E. A., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self-conscious emotion: insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology*, *85*, 594–604.

Beer, J. S., & Hughes, B. L. (2010). Neural systems of social comparison and the "above-average" effect. *NeuroImage*, *49*, 2671–2679.

Beer, J. S., John, O. P., Scabini, D., & Knight, R. T. (2006). Orbitofrontal cortex and social behavior: integrating self-monitoring and emotion-cognition interactions. *Journal of Cognitive Neuroscience*, *18*, 871–879.

Beer, J. S., Lombardo, M. V., & Bhanji, J. P. (2010). Roles of medial prefrontal cortex and orbitofrontal cortex in self-evaluation. *Journal of Cognitive Neuroscience*, *22*, 2108–2119.

Bem, D. J. (1967). Self-perception: an alternative interpretation of cognitive dissonance phenonmena. *Psychological Review*, *74*, 183–200.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.

Bhanji, J. P., & Beer, J. S. (2013). Dissociable neural modulation underlying lasting first impressions, changing your mind for the better, and changing it for the worse. *Journal of Neuroscience*, *33*, 9337–9344.

Blackwood, N. J., Bentall, R. P., Simmons, A., Murray, R. M., & Howard, R. J. (2003). Self-responsibility and the self-serving bias: an fMRI investigation of causal attributions. *NeuroImage*, *20*, 1076–1085.

Blumer, D., & Benson, D. F. (1975). Personality changes with frontal and temporal lobe lesions. In D. F. Benson & D. Blumer (Eds.). *Psychiatric Aspects of Neurologic Disease* (pp. 151–169). New York, NY: Grune & Stratton.

Bochner, S. (1994). Cross-cultural differences in the self concept: a test of Hofstede's individualism/collectivism distinction. *Journal of Cross-Cultural Psychology*, *25*, 273– 283.

Bond, M., & Cheung, T. (1983). College students' spontaneous self-concept: the effect of culture among respondents in Hong Kong, Japan, and the United States. *Journal of Cross-Cultural Psychology*, *14*, 153–171.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, *8*, 323–343.

Buunk, A. P., Cohen-Schotanus, J., & van Nek, R. (2007). Why and how people engage in social comparison while learning social skills in groups. *Group Dynamics: Theory, Research, and Practice*, *11*(3), 140–152.

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: the role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin*, *130*, 813–838.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*(5), 541–560. doi:10.1006/jesp.1997.1329

Cooley, C. H. (1902). *Human Nature and the Social Order*. New York, NY: Scribners.

Craik, F. I. M., Moroz, T. M., Moscovitch, M., Stuss, D. T., Wincour, G., Tulving, E., & Kapur, S. (1999). In search of the self: a positron emission tomography study. *Psychological Science*, *10*, 26–34.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *11*, 268–294.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*, 285–290.

Dunning, D. (1995). Trait importance and modifiability as factors influencing self-assessment and self-enhancement motives. *Personality and Social Psychology Bulletin*, *21*, 1297–1306.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: the role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*, 1082–1090.

Eastwick, P. W., Finkel, E. J., Krishnamurti, T., & Loewenstein, G. (2008). Mispredicting distress following romantic breakup: revealing the time course of affective forecasting error. *Journal of Experimental Social Psychology*, *44*, 800–807.

Eisenberg, N. (2010). Empathy-related responding: links with self-regulation, moral judgment, and moral behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature* (pp. 129–148). Washington, DC: American Psychological Association.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*, 203–210.

Flagan, T., & Beer, J.S. (2013). Three ways in which midline regions contribute to self-evaluation. *Frontiers in Human Neuroscience*, *7*, 1–12.

Fossati, P., Hevenor, S. J., Graham, S. J., Grady, C., Keightley, M. L., Craik, F., & Mayberg, H. (2003). In search of the emotional self: an FMRI study using positive and negative emotional words. *American Journal of Psychiatry*, *160*, 1938–1945.

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, *69*, 227–236.

Gilihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, *131*, 76–97.

Goethals, G. R., & Darley, J. (1977). Social comparison theory: an attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social Comparison Processes: Theoretical and Empirical Perspectives* (pp. 86–109). Washington, DC: Hemisphere.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: judgments of personality based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*, 379–398.

Gramzow, R. H., & Willard, G. (2006). Exaggerating current and past performance: motivated self-enhancement versus reconstructive memory. *Personality and Social Psychology Bulletin*, *32*, 1114–1125.

Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *23*, 443–451.

Heatherton, T. F., Macrae, C. N., & Kelley, W. (2004). What can the social brain science tell us about the self? *Current Directions in Psychological Science*, *13*, 190–193.

Hughes, B. L., & Beer, J. S. (2012a). Medial orbitofrontal cortex is associated with shifting decision thresholds in self-serving cognition. *NeuroImage*, *61*, 889–898.

Hughes, B. L., & Beer, J. S. (2012b).Orbitofrontal cortex and anterior cingulate cortex are modulated by motivated social cognition. *Cerebral Cortex*, *22*, 1372–1381.

Hughes, B. L., & Beer, J. S. (2013). Protecting the self: the effect of social-evaluative threat on neural representations of self. *Journal of Cognitive Neuroscience*, *25*, 613–622.

Ingram, R. E., Smith, T. W., & Brehm, S. S. (1983). Depression and information-processing: self-schemata and the encoding of self-reference information. *Journal of Personality and Social Psychology*, *45*, 412–420.

James, W. (1890). *The Principles of Psychology*. Cambridge, MA: Harvard University Press. (Originally published 1890.)

Jones, E. E., & Nisbett, R. E. (1971). *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior*. Morristown, NJ: General Learning Press.

Keenan, J. M., & Baillet, S. D. (1980). Memory for personally and socially significant events. In R. S. Nickerson (Ed.), *Attention and Performance* (Vol. 8, pp. 651–669). Hillsdale, NJ: Erlbaum.

Keenan, J. P., Wheeler, M. A., Gallup, G. G., & Pasucal-Leone, A. (2000). Self-recognition and the right prefrontal cortex. *Trends in Cognitive Sciences*, *4*, 338–344.

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785–794.

Kihlstrom, J. F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 17, pp. 1–47). New York, NY: Academic Press.

Kim, K., & Johnson, M. K. (2012). Extended self: medial prefrontal activity during transient association of self and objects. *Social Cognitive and Affective Neuroscience*, *7*(2), 199–207.

Kircher, T. T., Brammer, M., Bullmore, E., Simmons, A., Bartels, M., & David, A. S. (2002). The neural correlates of intentional and incidental self processing. *Neuropsychologia*, *40*, 683–692.

Kircher, T. T., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., . . . David, A. S. (2000). Towards a functional neuroanatomy of self processing: effects of faces and words. *Cognitive Brain Research*, *10*, 133–144.

Klayman, J., Soll, J. B., Gonzalez-Vallejo, C., & Barlas, S. (1999). Overconfidence: it depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, *79*, 216–247.

Klein, S. B., Babey, S. H., & Sherman, J. W. (1997). The functional independence of trait and behavioral self-knowledge: methodological considerations and new empirical findings. *Social Cognition*, *15*, 183–203.

Klein, S. B., & Kihlstrom, J. F. (1986). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General*, *115*(1), 26–38. doi:10.1037/0096-3445.115.1.26

Klein, S. B., Loftus, J., & Burton, H. A. (1989). Two self-reference effects: the importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-reference encoding. *Journal of Personality and Social Psychology*, *56*, 853–865.

Klein, S. B., Loftus, J., Trafton, J. G., & Fuhrman, R. W. (1992). Use of exemplars and abstractions in trait judgments: a model of trait knowledge about the self and others. *Journal of Personality and Social Psychology*, *63*, 739–753.

Knee, C., Patrick, H., & Lonsbary, C. (2003). Implicit theories of relationships: orientations toward evaluation and cultivation. *Personality and Social Psychology Review*, *7*(1), 41–55.

Kruglanski, A. W., & Mayseless, O. (1990). Classic and current social comparison research: expanding the perspective. *Psychological Bulletin*, *108*, 195–208.

Krusemark, E. A., Campbell, W.K., & Clementz, B. A. (2008). Attributions, deception, and event related potentials: an investigation of the self-serving bias. *Psychophysiology*, *45*, 511–515.

Kwan, V. S. Y., Barrios, V., Ganis, G., Gorman, J., Lange, C., Kumar, M., . . . Keenan, J. P. (2007). Assessing the neural correlates of self-enhancement bias: a transcranial magnetic stimulation study. *Experimental Brain Research*, *182*, 379–385.

Kwang, T., & Swann, W. B. (2010). Do people embrace praise even when they feel unworthy? A review of critical tests of self-enhancement versus self-verification. *Personality and Social Psychology Review*, *14*(3), 263–280.

Leary, M. R., Estrada, M. J., & Allen, A. (2009). The analogue-I and the analogue-Me: the avatars of the self. *Self and Identity*, *8*, 147–161.

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: the sociometer hypothesis. *Journal of Personality and Social Psychology*, *68*, 518–530.

Lemay, E. P. Jr., Clark, M. S., & Greenberg, A. (2010). What is beautiful is good because what is beautiful is desired: physical attractiveness stereotyping as projection of interpersonal goals. *Personality and Social Psychology Bulletin*, *36*(3), 339–353.

Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self-knowledge. *Cerebral Cortex*, *14*, 647–654.

Maki, R. H., & McCaul, K. D. (1985). The effects of self-reference versus other reference on the recall of traits and nouns. *Bulletin of the Psychonomic Society*, *23*, 169–172.

Malle, B. F. (2006). The actor-observer asymmetry in attribution: a (surprising) meta-analysis. *Psychological Bulletin*, *132*, 895–919.

Marcus, D. K., & Miller, R. S. (2003). Sex differences in judgments of physical attractiveness: a social relations analysis. *Personality and Social Psychology Bulletin*, *29*, 325–335.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, *35*, 63–78.

Markus, H., Hamill, R., & Sentis, K. P. (1987). Thinking fat: self-schemas for body weight and the processing of weight relevant information. *Journal of Applied Social Psychology*, *17*, 50–71.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Markus, H. [R.], & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Mead, G. H. (1934). *Mind, Self and Society*. Chicago, IL: University of Chicago Press.

Meier, B., Robinson, M., Carter, M., & Hinsz, V. (2010). Are sociable people more beautiful? A zero-acquaintance analysis of agreeableness, extraversion, and attractiveness. *Journal of Research in Personality*, *44*, 293–296.

Mitchell, J. P., Banaji, M., & Macrae, C. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306–1315.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*, 655–663.

Moran, J. M., Macrae, C. N., Heatherton, T. F., Wyland, C. L., & Kelley, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, *18*, 1586–1594.

Murray, S. L., & Holmes, J. G. (1997). A leap of faith? Positive illusions in romantic relationships. *Personality and Social Psychological Bulletin*, *23*, 586–597.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The benefits of positive illusions: idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology*, *70*, 79–98.

Murray, R. J., Schaer, J., & Debbane, M. (2012). Degrees of separation: a quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self- and other-reflection. *Neuroscience and Biobehavioral Reviews*, *36*, 1043–1059.

Mussweiler, T., & Rüter, K. (2003). What friends are for! The use of routine standards in social comparison. *Journal of Personality and Social Psychology*, *85*(3), 467–481.

Neff, L., & Karney, B. (2005). To know you is to love you: the implications of global adoration and specific accuracy for marital relationships. *Journal of Personality and Social Psychology*, *88*(3), 480–497.

Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs, NJ: Prentice Hall.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology*, *4*, 681–691.

Ochsner, K. N., Beer, J. S., Robertson, E. A., Cooper, J., Gabrieli, J. D. E., Kihlstrom, J. F., & D'Esposito, M. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, *28*, 797–814.

Parfit, D. (1984). *Reasons and Persons*. Oxford, UK: Oxford University Press.

Paulhus, D. L., Graf, P., & Van Selst, M. (1989). Attentional load increases the positivity of self-presentation. *Social Cognition*, *7*, 389–400.

Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D. (2009). Neural correlates of direct and reflected self-appraisals in adolescents and adults: when social perspective-taking informs self-perception. *Child Development*, *80*, 1016–1038.

Pinter, B., Green, J. D., Sedikides, C., & Gregg, A. P. (2011). Self-protective memory: separation/integration as a mechanism for mnemic neglect. *Social Cognition*, *29*, 612–624.

Plaks, J., Levy, S., & Dweck, C. (2009). Lay theories of personality: cornerstones of meaning in social cognition. *Social and Personality Psychology Compass*, *3*, 1069–1081.

Pronin, E., Lin, D., & Ross, L. (2002). The bias blind spot: perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*(3), 369–381.

Pronin, E., Olivola, C. Y., & Kennedy, K. A. (2008). Doing unto future selves as you would do unto others: psychological distance and decision making. *Personality and Social Psychology Bulletin*, *34*, 224–236.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: short-term benefits and long-term costs. *Journal of Personality and Social Psychology*, *80*, 340–352.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*, 677–688.

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., & Smith, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, *55*, 594–602.

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: food for thought. *Perspectives on Psychological Science*, *3*, 102–116.

Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, *450*, 102–105.

Srivastava, S., & Beer, J. S. (2005). How self-evaluations relate to being liked by others: integrating sociometer and attachment perspectives. *Journal of Personality and Social Psychology*, *89*, 966–977.

Story, A. (1998). Self-esteem and memory for favorable and unfavorable personality feedback. *Personality and Social Psychology Bulletin*, *24*, 51–64.

Suls, J., Lemos, K., & Stewart, H. L. (2002). Self-esteem, construal, and comparisons with the self, friends, and peers. *Journal of Personality and Social Psychology*, *87*, 252–261.

Swann, W. B., Pelham, B. W., & Krull, D. S. (1989) Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *Journal of Personality and Social Psychology*, *57*, 782–791.

Swann, W. B., Hixon, J. G., Stein-Seroussi, A., & Gilbert, D. T. (1990). The fleeting gleam of praise: cognitive processes underlying behavioral reactions to self-relevant feedback. *Journal of Personality and Social Psychology*, *59*, 17–26.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. *Psychological Bulletin*, *12*, 371–394.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Taylor, S. E., & Crocker, J. (1981). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social Cognition: the Ontario Symposium* (Vol. 1, pp. 89–134). Hillsdale, NJ: Erlbaum.

Taylor, S. E., & Koivumaki, J. H. (1976). The perception of self and others: acquaintanceship, affect, and actor-observer differences. *Journal of Personality and Social Psychology*, *33*, 403–408.

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: downward evaluation and upward contacts. *Psychological Review*, *96*, 569–575.

Taylor, S. E., Neter, E., & Wayment, H. A. (1995). Self-evaluation process. *Personality and Social Psychology Bulletin*, *21*, 1278–1287.

Tice, D. M. (1992). Self-presentation and self-concept change: the looking-glass self is also a magnifying glass. *Journal of Personality and Social Psychology*, *63*, 435–451.

Trope, Y. (1975). Seeking information about one's ability as a determinant of choice among tasks. *Journal of Personality and Social Psychology*, *16*, 116–123.

Trope, Y. (1980). Self-assessment, self-enhancement, and task preference. *Journal of Experimental Social Psychology*, *16*, 116–129.

Trope, Y. (1986). Self-enhancement and self-assessment in achievement behavior. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (pp. 350–378). New York, NY: Guilford Press.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*, 440–463.

Trope, Y., & Neter, E. (1994). Reconciling competing motives in self-evaluation: the role of self-control in feedback seeking. *Journal of Personality and Social Psychology*, *66*, 646–657.

Turk, D. J., Heatherton, T. F., Macrae, C. N., Kelley, W. M., & Gazzaniga, M. S. (2003). Out of contact, out of mind: the distributed nature of self. *Annals of the New York Academy of Sciences*, *1001*, 65–78.

Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social *cognition*: the role of mirror neurons. *Trends in Cognitive Sciences*, *11*, 153–157.

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences*, *7*, 65–69.

Wilson, T., & Gilbert, D. (2005). Affective forecasting: knowing what to want. *Current Directions in Psychological Science*, *14*, 131–134.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, *10*, 109–120.

CHAPTER 17

Motivation

EDDIE HARMON-JONES, TOM F. PRICE, AND CINDY HARMON-JONES

INTRODUCTION

Motivation underlies much perception, cognition, and behavior. As such, it is studied from many levels of analysis within psychology and neuroscience, ranging from the study of single cells and chemicals up to social relationships. In this chapter, we review research and theory on the psychology of motivation from the perspective of social emotive neuroscience, which is an integrative approach involving social/personality psychology, psychophysiology, affective neuroscience, and the study of emotions and motivation within humans.

Motivation can be defined simply as the urge or impulse to move or act. It varies in intensity from zero or no motivation to an extremely high level. In addition to this major dimension of motivational intensity, another major dimension in motivation is motivational direction, or the motivation to approach or go toward something as compared to the motivation to avoid or go away from something (E. Harmon-Jones, Harmon-Jones, & Price, 2013). This simple definition of motivation differs from the more commonly accepted one that motivation is "any internal process that energizes, directs, and sustains behavior" (Reeve, 2016, p. 31). Motivations certainly do energize, direct, and sustain behavior, but often motivations do not result in behavior. Consider the occasions individuals may be motivated to eat an entire chocolate cake but suppress that motivation because of fear of retribution from their stomach or others who wanted some of the same cake.

Motivation can arise from within the organism, and these internal causes can be more or less conscious to the individual. For instance, one can wake up on the motivated side of the bed and have lots of energy all day that could be applied to various tasks or external events. Motivations of this sort have received much less research attention than the other broad cause of motivation, external events. However, research on bipolar disorder has shed some light on this issue. When individuals are in manic phases, they are more motivated to approach various things, whereas when they are in depressive phases, they are less motivated to approach (Nusslock et al., 2011, 2012).

External events that trigger motivation often are seen or heard; that is, they are processed by the visual or auditory system. As such, they engage the orienting response because they attract attention. Moreover, these events often evoke approach or avoidance motivation and behavior (Thorndike, 1911) or, relatedly, emotion (Bradley, 2009; Maltzman, 1979). The processes of attending and orienting have been posited to stem from defensive or appetitive activations of motivational systems, which likely have evolved to sustain and protect the life of the individual (Bradley, 2009). Attention, emotion, and motivation, therefore, are inextricably linked. (See also Chapters 9, 15, and 18 in this volume; Nobre, 2017)

Emotion often is theorized to be a fundamental disposition to act or behave effectively to events that promote or threaten life (Frijda, 1986; Lang, 1985). The motivational tendencies associated with specific emotional states are situated in general systems of avoidance and approach, with avoidance processes preventing threats to well-being and approach processes promoting survival. Theorists have suggested that positivity judgments reflect approach motivation, negativity judgments reflect avoidance motivation, and arousal judgments index the intensity of activation or motivation (Bradley, 2009). Although this often may be the case, the relationship between motivational direction and emotional valence is not always that direct. For instance, anger, a negatively valenced emotion, often is linked with approach motivation (Carver & Harmon-Jones, 2009), a point elaborated later in this review (see the section titled "Asymmetric Frontal Cortical Activity and Anger").

Arousal often is posited to be an index of motivational intensity, with increased positive or negative arousal indicating more intense approach or avoidance motivation, respectively (Bradley & Lang, 2007). Although arousal likely does correspond to motivational intensity in many cases, it is not invariably associated with motivational intensity. For instance, humor and amusement are also emotional states associated with high arousal (Fredrickson & Branigan, 2005; Gable & Harmon-Jones, 2008), but these states do not necessarily urge action toward something, and so they are not associated with high levels of approach motivation. Similarly, caffeine and physical exercise increase subjective and physiological indices of arousal but not the motivation to approach or avoid (Gable & Harmon-Jones, 2013).

MOTIVATIONAL INTENSITY

In this section, we consider the broad motivational dimension of intensity and how it influences or relates to psychophysiological and cognitive responses.

Motivational Intensity and Psychophysiological Responses

Psychophysiological research has revealed several measures that are tightly related to motivational intensity. In particular, certain measures derived from the electroencephalogram (EEG) and from cardiovascular indices are associated with motivational intensity. Regarding the EEG, event-related potentials (ERPs) are derived from the EEG typically via signal averaging. That is, EEG activity in relation to some meaningful event (e.g., onset of a stimulus) is averaged over repeated trials of the same event type. For example, a researcher might average the EEG activity that occurs the 1,000 ms after the onset of an angry face. When the EEG is averaged over a sufficient number of trials (depending on stimulus complexity and other variables), peaks and valleys appear in the resulting signal, which is referred to as the averaged ERP. These peaks and valleys are positive and negative in relationship to the zero point electrically, and they also have a latency from the onset. For example, a P300 is a positive-going ERP about 300 ms after stimulus onset. Instead of having the positivity/negativity and latency form the basis of their name, some ERPs are labeled according to their psychological significance.

Event-Related Potentials

One ERP, the late positive potential (LPP), is sensitive to motivational intensity. It is not sensitive to motivational direction. The LPP starts approximately 300 ms after stimulus onset and occurs for several hundred milliseconds. (See Hajcak, Weinberg, MacNamara, & Foti, 2011.) LPP amplitudes are larger to both approach- and avoidance-related motivational stimuli relative to neutral stimuli. LPP amplitudes are also larger to more intense motivational stimuli (e.g., erotica) relative to less intense motivational stimuli (e.g., exciting sports scenes; Briggs & Martin, 2009). Relatedly, LPP amplitudes are larger to sexually explicit stimuli relative to slightly erotic stimuli (Prause, Steele, Staley, & Sabatinelli, 2014).

Individual differences in approach motivational strength relate to LPP amplitudes to approach-related stimuli (Gable & Poole, 2014). In this study (Gable & Poole, 2014), individual differences were measured with the behavioral inhibition sensitivity (BIS) and behavioral activation sensitivity (BAS) scales (Carver & White, 1994). Photographs were used to evoke anger and neutral affect; the anger-evoking photographs depicted anti-American acts, and they were shown to Americans who are patriotic. These anti-American anger-evoking photographs were expected to evoke approach motivation as these patriotic viewers may be motivated to confront the situations causing the anger. (We return to this point later in the section titled "Asymmetric Frontal Cortical Activity and Anger.") LPP amplitudes were larger to anger-evoking photographs than to neutral photographs. Moreover, greater trait approach motivation (BAS) was positively correlated to larger amplitudes to the anger-evoking photographs but not to the neutral photographs.

Again, however, LPP amplitudes are not specific to approach motivation. For example, photographs of threats and mutilations evoke larger LPP amplitudes than photographs of contamination and loss (Schupp et al., 2004). LPP amplitudes are larger to stimuli that evoke more intense motivation regardless of whether the motivation is avoidance or approach oriented. The LPP has been found to be generated by the parietal (Keil et al., 2002) and the occipitotemporal cortex (Sabatinelli, Lang, Keil, & Bradley, 2007).

Cardiovascular Responses

Cardiovascular responses, such as increases in systolic blood pressure and heart rate (HR), have been found to be associated with motivational intensity. One line of research that has provided much evidence on this point was derived from Brehm's (Brehm et al., 1983; Brehm & Self, 1989) theory of motivational intensity. According to this theory, motivational intensity is determined by potential motivation and perceived task difficulty, when it is known. Potential motivation sets the upper limit of how hard organisms are willing to work. Potential motivation is determined by the need for the particular outcome, the subjective value of the outcome, and the likelihood of the outcome being delivered or provided once successful task performance occurs. In this theory, these variables do not influence motivational intensity directly; they simply set an upper limit, and the perceived difficulty of obtaining the outcome determines motivational intensity. If an organism does not have a perception or belief about the difficulty of obtaining the outcome, motivational intensity will be determined by potential motivation. In other words, in situations in which task difficulty is unknown, motivational intensity will be determined by the variables that influence potential motivation—need state, subjective value of the outcome, and outcome expectancy.

Because organisms tend to conserve energy, they should not exert much effort when it is easy to obtain the desired outcome. But they will exert much more effort to obtain the desired outcome when it is moderately difficult to do so, assuming they value and need the outcome. However, if the organism perceives the outcome as impossible to obtain, it will exert no effort. In other words, motivational intensity is predicted to vary nonmonotonically (i.e., it looks like a shark's fin) with the perceived difficulty of obtaining the outcome. (See Ach, 1935; Hillgruber, 1912.) When potential motivation is relatively high, motivational intensity will increase for easy to moderately difficult tasks and then drop to a low level for impossible tasks. In contrast, when potential motivation is relatively low, motivational intensity might be somewhat high for easy tasks but drop to a low level for moderately difficult and impossible tasks. (For a review, see Gendolla, & Wright, 2012.)

Research derived from this theory has used measures of cardiovascular system activity as an index of motivational intensity. The idea that cardiovascular system activity would relate to motivational intensity is based on two assumptions: (1) One of the main functions of the cardiovascular system is to sustain behavior (Papillo & Shapiro, 1990); and (2) effort influences cardiovascular responses via activation of certain sympathetic nervous system mechanisms (Elliott, 1969; Obrist, 1976). According to some theorists (Wright, 1996), heart contraction force (i.e., cardiac contractility, often measured as pre-ejection period) and variables influenced by it, particularly systolic blood pressure (pressure at the peak of a pulse), are some of the best indicators of these sympathetic mechanisms. HR—the pace of heart contraction—is an indicator of these sympathetic mechanisms, but it is also influenced by parasympathetic activity that occasionally can obscure or reverse the sympathetic influences.

Several experiments have tested this theory. For example, Wright (1984) told participants they could avoid an electric shock by successfully performing a motor task, which was designed to be difficult or easy (a dynamometer grip or a toggle switch, respectively). Participants in a control condition had no behavior available to avoid the electric shock. Results indicated that HR and finger pulse volume were greatest in the difficult task condition (immediately prior to the task for the task conditions, or while control condition participants paused for further instructions). These results provide support for the idea that motivational intensity is influenced by perceived task difficulty even in an aversive context.

Other experiments have provided evidence consistent with this theory. For instance, Richter, Friedrich, and Gendolla (2008) had participants perform a memory task that varied in task difficulty while pre-ejection period, HR, and blood pressure were recorded. Results indicated that pre-ejection period and systolic blood pressure increased as task difficulty increased until the fourth level of difficulty. At this point, when task success was impossible, pre-ejection period and systolic blood pressure were low. These results provide further support for the idea that motivational intensity is influenced by perceived task difficulty in a nonmonotonic fashion and that motivational intensity is associated with beta-adrenergic impacts on the heart.

Influence of Motivational Intensity on Cognitive Processes

Motivationally intense states narrow cognitive scope; that is, they reduce the breadth of cognitive expansiveness. This narrowing of cognitive scope can occur at perceptual, attentional, or conceptual levels. (See review by Gable & Harmon-Jones, 2010d.) Motivationally intense states may narrow cognitive scope because a narrowed focus on the desired goal, whether appetitive or aversive, may assist in ultimately obtaining the goal.

Research beginning in the 1970s had suggested that positive affect broadens cognitive scope. (See review by Fredrickson, 2001.) However, this research had confounded affective valence with motivational intensity, such that only those positive affective states low in motivational intensity (e.g., amusement) were compared to those negative affective states high in motivational intensity (e.g., fear). To remove this confound, experiments were conducted in which positive affective states low versus high in motivational intensity were compared on their influence on cognitive scope.

In one of the first of such tests, low approach-motivated positive affect was induced with a humorous video clip similar to manipulations in past research on positive affect and broadening (Gable & Harmon-Jones, 2008). In comparison, high approach-motivated positive affect was induced with a video clip of delicious-looking desserts. Both videos caused participants to report feeling equal levels of general positive affect (e.g., happiness), but the video of desserts caused participants to report feeling more desire whereas the humorous clip caused participants to feel more amusement. To measure cognitive scope, the local-global task of Kimchi and Palmer (1982) was used. In this task, on each trial, three figures are displayed. The standard figure is situated at the top, and two comparison figures are situated below it. One of the comparison figures has local elements that are similar to the local elements of the standard; the other comparison figure has a global configuration that is similar to the global configuration of the standard. Participants were asked to give their "first and most immediate impression" of which of the two comparison figures was most similar to the standard figure. From their responses, it can be determined whether the perceiver had a more local (narrow; choice of comparison with local similarity) or global (broad; choice of comparison with global similarity) attentional scope. Results indicated that the video that evoked high motivationally intense positive affect caused less broadening than the video that evoked low motivationally intense positive affect.

In follow-up experiments, a neutral affect comparison condition was included, and attentional scope was measured using the Navon (1977) task. In this task, participants are shown compound stimulus letters, and they are asked to identify one of two letters (e.g., "H" or "T") on each trial. Each compound stimulus letter contains only one of the two letters, and it is displayed at the local (small) or global (large) level. The time it takes the participant to indicate correctly which of the two target letters is presented (which is at the local or global level) serves as the dependent variable. In theory, a broader attentional scope could be associated with faster detection of global letters or slower detection of local letters, whereas a narrower attentional scope could be associated with faster detection of local letters or slower detection of global letters. Replicating much past research, neutral stimuli caused broadened attention (Navon, 1977). More important, the appetitive stimuli caused more narrowed attention. Additional studies revealed that individuals who scored high in trait approach motivation (on the BAS questionnaire of Carver & White, 1994) had even more narrowed attention following appetitive stimuli. Experiments also revealed that increasing the approach motivational intensity of the stimuli (by giving participants the expectation that they would eat desserts displayed in the photographs) caused even more narrowed attention following appetitive stimuli (Gable & Harmon-Jones, 2008, Studies 1–4). In a conceptually similar manner, research has found that individuals motivated to consume alcohol had more narrowed attention following alcohol-related photographs (Hicks, Friedman, Gable, & Davis, 2012).

In this line of research, a variety of inductions of positive affect and measures of cognitive scope have been used. In two studies (Gable & Harmon-Jones, 2010b), cognitive scope was assessed with recognition memory for neutral words presented in the center versus the periphery of the viewing space. Results indicated that participants had better recognition memory for words presented in the center of the viewing space when they were in a high approach-motivated positive state as compared to a neutral state. In contrast, participants had better recognition memory for words presented in the periphery of viewing space when they were in a low approach-motivated positive state as compared to a neutral state. In one of the these experiments, low versus high approach positive affect was manipulated in a monetary incentive delay task (Knutson & Wimmer, 2007). In the other experiment, photographs were used to evoke high approach-motivated positive affect.

Manipulated body postures also influence approach motivational intensity (E. Harmon-Jones, Gable, & Price, 2011; Price, Dieckman, & Harmon-Jones, 2012). Based on these findings, in another set of studies (Price & Harmon-Jones, 2010b), high versus low approach-motivated positive affect was manipulated by having participants recline backward (low approach positive state), sit upright (moderate approach positive state), or lean forward (high approach positive state). In each body position, participants were told to configure their faces in a manner to cause a smile. Participants were informed we were examining how these various body positions influence brain activity. Cognitive scope was measured using a cognitive categorization task developed by Isen and Daubman (1984). This task presents participants with specific categories (e.g., vehicle) and typical (e.g., car) and atypical (e.g., camel) examples of these categories. Participants indicate how well example items fit into specific categories. Past research has demonstrated that participants in a positive state induced via a free gift rated the categorization examples more inclusively, that is, they were more likely to indicate that "camel" fits with the category "vehicle." Our results revealed that only the low approach positive state replicated this effect. The high approach positive state caused less inclusive categorization.

To test whether psychophysiological variables related to approach motivation and its intensity would relate to the cognitive narrowing that occurs as a result of approach motivation, studies have measured psychophysiological responses to the approach motivation primes and then tested whether these responses correlated with a more narrow cognitive scope. In one study (E. Harmon-Jones & Gable, 2009), asymmetric frontal cortical activity was measured using EEG alpha power. In a within-subjects design, participants were shown neutral or dessert pictures prior to responding to an attentional scope task (the Navon task described earlier). Results indicated that greater relative left frontal activity to appetitive photographs was correlated with more narrowing of attention immediately following those photographs. In another study (Gable & Harmon-Jones, 2010c), a different neural measure was used, the late positive potential of the ERP, which has been found to be a measure of motivational intensity. In this within-subjects design, participants also were shown neutral or dessert pictures prior to responding to an attentional scope task (the Navon task). Results indicated that larger late positive potentials to appetitive photographs were correlated with more narrowing of attention immediately following those photographs.

This research concerned positive affective states that differ in approach motivational intensity. The conceptual analysis—that increased motivational intensity will lead to more narrowed cognitive scope—also applies to negative affective states that differ in motivational intensity. Dating back to Easterbrook's (1959) classic review, much research has revealed that negative affect leads to a narrowing of cognitive scope, but this research has examined negative affective states that were high in motivational intensity (e.g., fear). Much less research has examined negative affective states that are low in motivational intensity, such as some forms of sadness and depression. However, consistent with the conceptual analysis offered here, some research has found that depression is associated with a broadened cognitive scope (von Hecker & Meiser, 2005).

To examine the effect of the motivational intensity of distinct negative affective states on cognitive scope, experiments have been conducted in which disgust, sadness, and neutral affect were compared in their influence on cognitive scope (Gable & Harmon-Jones, 2010a). These affective states were evoked using photographs, and attentional scope was measured with the Navon task. Results indicated that compared to neutral affect, disgust caused more self-reported arousal and sadness caused a similar amount of self-reported arousal. More important, compared to neutral affect, disgust caused a narrowing of attention and sadness caused a broadening of attention. Additional experiments have found that anger, an intense affective state associated with approach motivation (see the section titled "Asymmetric Frontal Cortical Activity and Anger"), causes a narrowing of attention (Gable, Poole, & Harmon-Jones, 2015).

Motivational intensity often is correlated with arousal. That is, motivationally intense states often are arousing states (both subjectively and physiologically). But motivational intensity can be separated from arousal. For example, as noted earlier, when individuals feel amused, they feel aroused but not motivationally intense or engaged, because amusement does not urge action toward something (E. Harmon-Jones et al., 2013). Research has indicated that this arousing state of amusement causes more attentional broadening than neutral states (Fredrickson & Branigan, 2005; Gable & Harmon-Jones, 2008), suggesting that arousal is not the key variable in influencing cognitive scope.

In addition, arousal can be created from situations such as physical exercise that do not necessarily influence motivational intensity (Gable & Harmon-Jones, 2013). To test the effects of nonmotivation arousal on cognitive scope, one experiment had participants pedal a stationary bike or not while completing the appetitive versus neutral picture/attentional scope task described previously. Participants in the pedaling condition had higher HRs than individuals who did not pedal, which indicated that they were aroused. More important, this arousal manipulation had no effect on attentional scope, even though the appetitive picture manipulation did. These results suggest that motivational intensity, rather than arousal per se, influences the narrowing of cognitive scope.

MOTIVATIONAL DIRECTION

In this section, we consider the broad motivational dimension of direction and how it influences or relates to psychophysiological and cognitive responses.

Motivational Direction and Psychophysiological Responses

Several conceptual models postulate that approach motivation and reward responses are associated with a positive affective system, whereas avoidance motivation and punishment responses are associated with a negative affective system. This postulate, however, has been challenged with research that has revealed that anger, a negative emotion, is associated with approach motivation. Because motivation and affect (or emotion) are so interlinked, next we consider research on affect. One neural pattern of activity that has received extensive tests in relationship to motivational direction is asymmetric frontal cortical activity.

Asymmetric Frontal Cortical Activity

The reciprocal/asymmetric involvement of prefrontal cortical regions in positive and negative emotions as well as approach and withdrawal motivations, respectively, was suggested over 80 years ago by observations of individuals who had suffered brain damage to the right and left anterior cortex (Goldstein, 1939). In the 1930s, it was observed individuals who had suffered damage to the left prefrontal cortex (PFC) were more likely to develop depression. Research employing the Wada test supported these early observations. The Wada test involves injecting a barbiturate derivative, sodium amytal, into one of the internal carotid arteries, which suppresses the activity of one hemisphere. Injecting amytal in the right artery produced euphoria; injecting amytal in the left artery produced depressed affect (Alema, Rosadini, & Rossi, 1961; Perria, Rosadini, & Rossi, 1961; Rossi & Rosadini, 1967; Terzian & Cecotto, 1959). These effects suggested that the hemispheres of the brain are reciprocal; it is possible to release one hemisphere from the contralateral inhibitory influences of the other. From this perspective, left hemispheric activation, when not inhibited by the right hemisphere, induced euphoria; an uninhibited right hemisphere induced depression.

Later studies confirmed these results, finding that patients who had suffered left hemispheric infarcts or lesions tended to show symptoms indicative of depression (Black, 1975; Gasparrini, Satz, Heilman, & Coolidge, 1978; Gainotti, 1972; Robinson & Price, 1982), whereas patients who had suffered right hemispheric infarcts or lesions tended to show symptoms indicative of mania (Gainotti, 1972; Robinson & Price, 1982; Sackeim et al., 1982). Other research has suggested that asymmetries connected to appetitive and avoidant behaviors occur in nonhuman animals, in species ranging from great apes and reptiles (Deckel, Lillaney, Ronan, & Summers, 1998; Hopkins, Bennett, Bales, Lee, & Ward, 1993), chicks (Güntürkün et al., 2000), amphibians (Rogers, 2002), to spiders (Ades & Ramires, 2002).

More recent research has indicated that in humans, these asymmetric activation patterns related to affect are often specific to the frontal cortex. This research often uses right versus left frontal cortical asymmetric activation as a dependent variable, usually examined with EEG recordings. Frontal cortical asymmetry is calculated by comparing cortical activation levels between analogous areas on the left and right sides. Difference scores often are used in this research; their use is consistent with the lesion and amytal research just described, which suggests that asymmetry may be the crucial variable, with one hemisphere inhibiting the opposing one.

Much of this evidence has been obtained with measures of alpha frequency band activity derived from the EEG brain activity. Research has indicated that alpha power is inversely related to regional brain activity using behavioral tasks (Davidson, Chapman, Chapman, & Henriques, 1990) and hemodynamic measures (Cook, O'Hara, Uijtdehaage, Mandelkern, & Leuchter, 1998). Functional magnetic resonance imaging results (Berkman & Lieberman, 2010) and the source localization of EEG signals (Pizzagalli, Sherwood, Henriques, & Davidson, 2005) obtained in emotive frontal asymmetry studies converge in indicating that the dorsolateral PFC is responsible for these effects. These findings also are corroborated by studies of transcranial magnetic stimulation, discussed later (Schutter, 2009; Schutter, van Honk, d'Alfonso, Postma, & de Haan, 2001).

Resting Frontal Cortical Asymmetry and Trait Affective Styles

Studies have indicated that depression relates to resting frontal asymmetric activity. In particular, depressed individuals often evince relatively less left than right frontal cortical activity (Jacobs & Snyder, 1996; Schaffer, Davidson, & Saron, 1983), even when they are in remission (Henriques & Davidson, 1990). Other research has revealed that healthy individuals high in trait positive affect evince greater left than right frontal cortical activity, whereas individuals high in trait negative affect evince greater right than left frontal activity (Tomarken, Davidson, Wheeler, & Doss, 1992).

Subsequent studies indicated that trait approach motivation is related to greater left than right frontal activity at a resting baseline (Amodio, Master, Yee, & Taylor, 2008; Harmon-Jones & Allen, 1997; Sutton & Davidson, 1997). More recently, relative right frontal activity also has been associated with neuroticism and anxiety (Uusberg et al., 2015). In a study by Uusberg et al. (2015), participants completed a neuroticism scale and were asked to view a stimulus individual making direct eye contact, avoiding eye contact, or closing the eyes while EEG was recorded. Results indicated that individuals high in neuroticism viewing a face with a direct gaze had greater relative right frontal activity than all of the other conditions.

Thus, studies suggested that asymmetric frontal cortical activity is not necessarily associated with affective valence but rather with the motivational direction of affects. Approach and avoidance motivation, however, are mostly associated with positive and negative affect, respectively (Carver & White, 1994). Consequently, the interpretation was clouded. Similarly, the finding that promotion regulatory foci versus prevention foci are associated with greater relative left than right frontal activation at baseline (Amodio, Shah, Sigelman, Brazy, & Harmon-Jones, 2004) could be understood from either the affective valence or motivational direction view. Past research, therefore, had essentially confounded motivational direction with emotional valence. From these findings, researchers had interpreted that relatively greater right than left frontal cortical activity reflected greater withdrawal motivation and negative affect, whereas relatively greater left than right frontal cortical activity reflected greater approach motivation and positive affect. These claims fit well into dominant emotion theories associating negative affect with withdrawal motivation and positive affect with approach motivation (Lang, 1995; Watson, 2000).

Asymmetric Frontal Activity and State Affect

Research also has indicated that asymmetric frontal cortical activity is associated with state emotional responses. Davidson and Fox (1982) found that film clips of an actress generating a happy facial expression as compared to a sad facial expression caused 10-month-old infants to exhibit increased left frontal activation. Examining asymmetry to affective pictures has produced mixed results, as reviewed below. Some research, however, has suggested that pictures associated with basic motivational impulses, such as erotica, can elicit greater relative left frontal cortical activity relative to control stimuli (Schöne et al., 2016). In addition, relative right frontal cortical activity has been associated with induced withdrawal-oriented emotional states, such as fear and disgust (Davidson et al., 1990; Jones & Fox, 1992), and empathic distress (Tullet et al., 2012). More recently, relative right frontal cortical activity has been associated with nostalgia (Tullet et al., 2015).

Some positive affects are higher in approach motivation, whereas others are lower in approach motivation. An important question, however, remained. Namely, do positive affects of any approach motivational intensity (i.e., high or low) cause increases in relative left frontal activation? E. Harmon-Jones, Harmon-Jones, Fearn, Sigelman, and Johnson (2008) addressed this issue by manipulating approach motivation and positive affect independently while frontal EEG was measured. Participants wrote about positive/low-approach and positive/high-approach events. In each case, they reported elevated positive affect. Only the positive/high-approach manipulation, however, produced an increase in relative left-frontal activity. These results support the notion that it is not the positive valence per se but the approach motivational aspects of some forms of positive affect that cause greater relative left versus right frontal cortical activation, measured with EEG.

Asymmetric Frontal Cortical Activity and Anger

Furthermore, experiments have suggested that approach motivation and positive affect are not perfectly connected with each other (E. Harmon-Jones et al., 2008). Research with anger offers more convincing evidence for the dissociation of motivational direction and affective valence (e.g., positive affect equals approach motivation). Anger evokes behavioral tendencies of approach but is nevertheless a negatively valenced emotion (e.g., Darwin, 1872; Ekman & Friesen, 1975; Plutchik, 1980; Young, 1943). Anger is often associated with offensive aggression and attack (e.g., Berkowitz, 1993; Blanchard & Blanchard, 1984; Lagerspetz, 1969). Defensive aggression is often associated with fear and can be distinguished from offensive aggression. Other research has suggested that anger is associated with approach motivation (e.g., Amodio & Harmon-Jones, 2011; Izard, 1991; Lewis, Alessandri, & Sullivan, 1990; Lewis, Sullivan, Ramsay, & Alessandri, 1992). More recent studies have examined whether BAS relates to anger-related responses. Several studies have indicated that BAS, trait behavioral approach sensitivity, measured with Carver and White's (1994) scale, positively relates to state and trait anger (Carver, 2004; E. Harmon-Jones, 2003; Smits & Kuppens, 2005). The emotion of anger, therefore, provides a critical test. Studies on anger have helped to disentangle interpretations of frontal cortical asymmetry and affective valence versus motivational direction.

To test this notion, Harmon-Jones and Allen (1998) measured trait anger using the Buss and Perry (1992) questionnaire. These researchers then assessed asymmetric frontal activity during a resting-baseline EEG recording. In this study of adolescents, trait anger was associated with increased relative left frontal activity and decreased relative right frontal activity. Asymmetric activity in other brain regions did not relate with anger. The specificity of anger to frontal asymmetries and not to other brain asymmetries has been observed in each of the reviewed studies on anger. Further, these results have been replicated in a study revealing that anger is not evaluated as a positive feeling (E. Harmon-Jones, 2004) and in other studies (Hewig, Hagemann, Seifert, Naumann, & Bartussek, 2004; Rybak, Crayton, Young, Herba, & Konopka, 2006). Other research has manipulated asymmetrical frontal cortical activity using repetitive transcranial magnetic stimulation. These studies have found that the electrical disruption of the right PFC using repetitive transcranial magnetic stimulation increased approach responses to angry faces compared with disruptions of the left PFC (d'Alfonso et al., 2000; van Honk & Schutter, 2006).

Researchers also have manipulated anger at the state level to test the motivational directional model of frontal cortical asymmetry. E. Harmon-Jones and Sigelman (2001) found that individuals who were insulted evinced greater relative left than right frontal activity compared to individuals who were not insulted. Additional analyses indicated that, within the insult condition, self-reported anger and aggression were correlated positively with relative left frontal activity. These correlations were nonsignificant in the no-insult condition. These results were conceptually replicated and extended by showing that social rejection associated with anger and jealousy increases relative left frontal activity (E. Harmon-Jones, Peterson, & Harris, 2009). These patterns of approach motivation and state anger have been conceptually replicated in other laboratories as well (e.g., Jensen-Campbell, Knack, Waldrip, & Campbell, 2007; Verona, Sadeh, & Curtin, 2009).

Considered as a whole, these studies have suggested that approach versus withdrawal motivation is associated with relative left versus relative right frontal cortical activity, respectively. Moreover, they provide evidence that approach versus withdrawal motivation is not inevitably associated with positive versus negative affect, as several previous conceptual models had assumed.

Reward Positivity

Research has discovered an ERP that is sensitive to reward delivery, the Reward Positivity (RewP). The RewP is a positive-going wave that is highest in amplitude in response to rewarding outcomes, such as monetary gains compared to losses (Hajcak, Moser, Holroyd, & Simons, 2006; Proudfit, 2015). Previously, this ERP was conceived of as a negative-going wave that was greater for losses than gains. As such, this wave has been named the feedback-related negativity or feedback negativity (Hajcak et al., 2006), feedback error-related negativity (Holroyd, Hajcak, & Larsen, 2006), or medial frontal negativity (Gehring & Willoughby, 2002). Proudfit (2015), however, posited that the negative-going wave that appears greater for losses than gains actually reflects a positive-going wave that is greater for gains than losses. This interpretation is consistent with evidence indicating that both losing and breaking even always evoke a negativity whereas winning evokes a positivity (Kujawa, Smith, Luhmann, & Hajcak, 2013). That is, the RewP is largest in amplitude when feedback indicates that an action led to a reward, and it is smallest when feedback indicates that an action led to a loss (Hajcak et al., 2006) or absence of reward (Bellebaum, Polezzi, & Daum, 2010). The negativity to nonreward or punishment feedback is likely a "baseline" response, whereas winning (or rewarding feedback) evokes a positivity.

The RewP has been associated with multiple neural regions, including the ventral striatum (Carlson, Foti, Harmon-Jones, & Proudfit, 2015), midcingulate and midfrontal cortices (Becker, Nitsch, Miltner, & Straube, 2014), and anterior cingulate cortex

(Gehring & Willoughby, 2002; Hauser et al., 2014), and to a reward circuit involving the ventral striatum, medial PFC/anterior cingulate cortex, orbitofrontal cortex, amygdala, and caudate (Carlson, Foti, Mujica-Parodi, Harmon-Jones, & Hajcak, 2011).

The RewP has been found to be associated with motivational direction. In particular, withdrawal-related states and traits are associated with smaller RewP responses. For example, depressive tendencies are associated with smaller RewP responses (Foti & Hajcak, 2009; Foti, Kotov, Klein, & Hajcak, 2011). In addition, trait anxiety is associated with smaller RewP responses when feedback is unambiguous (Gu, Ge, Jiang, & Luo, 2010; Gu, Huang, & Luo, 2010). Moreover, trait negative affect is associated with smaller RewP responses (Santesso et al., 2012). Relatedly, a greater propensity to react with sadness during an experiment was associated with a smaller RewP (Foti & Hajcak., 2010). In contrast, approach-related states and traits are associated with larger RewP responses. For example, extraversion is associated with larger RewP responses (Cooper, Duke, Pickering, & Smillie, 2014; Smillie, Cooper, & Pickering, 2011). Moreover, trait approach motivation is associated with larger RewP responses (Bress & Hajcak, 2013). Other research has revealed that the RewP amplitude positively correlates with how much individuals like the reward stimuli, and this correlation is greater following a state induction of anger (Angus, Kemkes, Schutter, & Harmon-Jones, 2015). These results suggest that anger, a negative affect associated with approach motivation, increases the RewP particularly for individuals who like the reward stimuli.

Although the RewP is associated with motivational direction, it also is associated with the intensity of the motivational direction (as are most variables). One recent study illustrated this point by examining perceived control. As reviewed earlier, Brehm's theory of motivation intensity (Brehm & Self, 1989) posits that motivational intensity increases up to the point where success is perceived as impossible. When success is perceived as impossible or uncontrollable, then motivational intensity should be low. Thus, if persons believe they have no control over outcomes, they should have smaller RewP responses. In the study testing this hypothesis (Mühlberger, Angus, Jonas, Harmon-Jones, & Harmon-Jones, 2017), participants were told that the rewards would appear randomly (low perceived control condition), or they were told that they could learn a mouse-click rule to gain rewards (high perceived control condition). In both conditions, participants engaged in exactly the same behavior and viewed the same task stimuli. The only difference between the two conditions was psychological meaning of the task—the perception of control. Results revealed that participants in the high perceived control condition had larger RewP responses than participants in the low perceived control condition. Taken together with previous results, these results suggest that approach motivational intensity increases the RewP response.

Startle Eyeblink Modulation

The startle eyeblink reflex has been shown to involve the central nucleus of the amygdala (Davis, 2006) and to be modulated by the emotive significance of stimuli (Bradley, Codispoti, Cuthbert, & Lang, 2001; Lang, Bradley, & Cuthbert, 1990; Vrana, Spence, & Lang, 1988). This reflex is a component of the full startle response occurring to sudden aversive and unanticipated events. In laboratory settings, the startle response is elicited by presenting participants with loud (100 db) bursts of white noise with instantaneous rise times (Blumenthal et al., 2005).

When presented, the muscle around the eye, the orbicularis oculi, contracts. This response occurs to protect the eye from potential harm.

In laboratory experiments, startle probes often are presented during the viewing of affective pictures (as well as neutral comparison pictures). The magnitude of the startle eyeblink is modulated depending on if the picture is associated with an appetitive or avoidant motivational state. Startle responses are diminished when viewing appetitive motivational pictures (compared to when viewing neutral pictures). Startle responses are potentiated when viewing avoidant motivational pictures (compared to when viewing neutral pictures). These effects often are explained with the response matching hypothesis, which suggests that the startle response is initiated by aversive stimuli (the startle probe) and therefore is a defensive response. The eyeblink magnitude is determined by whether the affective picture matches or does not match the aversive motivation elicited by the startle stimulus. If an aversive picture eliciting avoidance motivation is presented during the startle stimulus, then the aversive content of both the stimulus and the picture match, increasing the startle response. If an appetitive picture eliciting approach motivation is presented during the stimulus, then the aversive content of the startle probe and the picture mismatch, decreasing the startle response. Based on this hypothesis, more appetitive responses to stimuli are thought to be associated with smaller startle responses.

In line with the hypothesis that attenuated startle responses during the viewing of appetitive pictures reflect approach motivational impulses, individuals high in trait behavioral approach evince smaller startle response during arousing positive pictures (Hawk & Kowmas, 2003). In addition, individuals high in trait approach-oriented emotions, such as enjoyment, anger, and surprise, evince smaller startle response during arousing positive pictures (Amodio & Harmon-Jones, 2011). Similar effects have been observed in individuals higher in sensitivity to rewards (Aluja et al., 2014). Finally, positive images associated with more approach motivation (e.g., erotica) elicit smaller startle responses than positive images associated with less approach motivation (e.g., sailboats; Gard, Gard, Mehta, Kring, & Patrick, 2007). Thus, the emotive modulated startle reflex is sensitive to both motivational direction and motivational intensity.

Motivational Direction and Cognitive Responses

Motivational direction also has been found to be related to various cognitive responses. Much research has found that negative emotions associated with withdrawal (e.g., fear, anxiety) lead individuals to selectively attend to threatening information (e.g., MacLeod, 1999), and positive emotions associated with approach (e.g., happy, excitement) lead individuals to selectively attend to rewarding information (e.g., Tamir & Robinson, 2007).

More recent research on selective attention to threats and rewards has examined anger to determine whether affective valence or motivational direction best accounts for the previous results. These studies on anger have found that situational manipulations of anger cause an increase in visual attention to rewarding but not threatening information (Ford et al., 2010). Moreover, trait anger has been found to be associated with selective attention to rewarding but not threatening information, whereas trait anxiety has been found to be associated with selective attention to threatening but not rewarding information (Ford, Tamir, Gagnon, Taylor, & Brunyé., 2012).

Other research has linked motivational direction with memory for emotional events. For example, Gomez and Gomez (2002) found that individual differences in approach and avoidance motivation were associated with memory for emotional events. That is, individual differences in approach motivation were positively correlated with the number of positive words correctly identified in a word recognition task and the number of positive words recalled in a free recall task. In contrast, individual differences in avoidance motivation were positively correlated with number of negative words correctly identified in a word recognition task and the number of negative words recalled in a free recall task.

EMBODIMENT OF MOTIVATIONAL DIRECTION AND INTENSITY

The notion that physical postures and facial expressions are tightly connected with emotional experiences stems from theorizing by William James (1890) and Charles Darwin (1872). Such theories gave rise to facial feedback theories, suggesting that the mere manipulation of facial muscles can influence emotional reactions to stimuli (Laird, 1974), and "free-floating" emotional feelings (Duclos et al., 1989). The notion that such manipulations also may influence motivational approach and avoidance has been theorized (Zajonc et al., 1982). Although much work concerns the motivational intensity of embodiment, some research has focused on motivational direction and needs to be considered. (See also Matheson & Barsalou, 2017.)

Facial Expressions

Some early research has tested if manipulated facial expressions influence asymmetric cortical activity associated with motivational intensity. Ekman and Davidson (1993) tested if genuine smiles with zygomatic major (cheek) and orbicularis oculi muscle contractions increase relative left frontal cortical activity more so than less genuine smiles with only zygomatic major muscle activity. Genuine smiles relative to other forms of smiles have been demonstrated to increase an individual's motivation to receive social rewards and therefore may elicit higher approach motivation than less genuine smiles (Heerey, 2014). Ekman and Davidson (1993) covertly instructed participants to move individual muscles to create emotional facial expressions consistent with facial feedback techniques. Results indicated that, indeed, genuine smiles elicited greater relative left frontal activity than less genuine smiles. This experiment is an example of approach motivational intensity, as both forms of the manipulation are approach-motivated.

Additional experiments have investigated manipulated facial expression differing in motivational direction. Coan, Allen, and Harmon-Jones (2001) manipulated emotional facial expressions of joy, anger, fear, and sadness while asymmetric cortical activity was recorded. Results indicated that facial expressions high in approach motivation, anger and joy, increased relative left frontal activity. Expressions less associated with approach motivation, fear and disgust, reduced relative left frontal activity.

More recent experiments have again investigated manipulated positive emotional facial expressions varying in approach motivational intensity. Price, Hortensius, and Harmon-Jones (2013) examined facial expressions of determination and satisfaction. Determination has been associated with heightened, pre-goal approach motivation (C. Harmon-Jones, Schmeichel, Mennitt, & Harmon-Jones, 2011), while satisfaction has been associated with lower, post-goal approach motivation, consistent with research demonstrating that positive emotions can vary in approach motivational intensity (Gable & Harmon-Jones, 2008). Price et al. (2013) explicitly instructed participants to make determination, satisfaction, or neutral facial expressions such that anyone would be able to recognize them. This method differs from typical covert muscle-by-muscle movement instructions. Facial expressions of determination, however, are morphologically similar to facial expressions of anger; participants often confuse the two expressions (C. Harmon-Jones, Schmeichel, Mennitt, et al., 2011). Thus, muscle-by-muscle instructions were not implemented in Price et al. (2013) to prevent participants from simply creating anger facial expressions.

Results from this experiment indicated that participants who created determination facial expressions evinced greater relative left frontal activity relative to a resting baseline. Satisfaction and neutral expressions, however, did not increase relative left frontal cortical activity relative to baseline. While maintaining the assigned facial expressions, furthermore, participants completed a motivational persistence measure that involved completing several geometric puzzles (Glass & Singer, 1972), some of which were unsolvable. The persistence metric involved the number of attempts participants made at solving the insolvable puzzles. Results indicated that, within the determination condition only, relative left frontal activity was positively related to the number of attempts at the insolvable puzzles. These results suggested that relative left frontal cortical activity elicited by determination is associated with more persistence behavior.

Whole-Body Postures

Some early research considered the role of physical postures on motivational direction. Riskin and Gotay (1982) designed two postures intended to influence motivation. They utilized prior accounts of expressions of depression to create a “depressed” posture. Their depressed posture had participants’ torsos bent forward from the waist, their chests and necks dropped downward, and their heads and necks pushed down and forward so that participants were hunched and stooped over. A second posture was also created, similar to the physical expression of pride (Tracey & Robins, 2007), which has been associated with heightened approach (Williams & DeSteno, 2008). This posture had participants’ shoulders pushed backward and slightly raised, with the spine straightened such that the back was upright and erect in an expansive position. Participants’ heads were slightly raised at the chin, looking forward and slightly upward. Participants held one of the two postures for 8 min, ostensibly to collect physiological data for a “biofeedback task.” Next, participants returned to a normal posture and participated in a second task on “spatial thinking,” which was the motivational persistence measure just described. Results indicated that the “depressed” posture compared to the expansive posture led participants to have less persistence.

These findings have been expanded. Nair et al. (2015) had participants complete a social stress test while in either a slumped or an expansive posture. The social stress test involves giving a speech for a supposed dream job while being video recorded. Results replicated those of Riskind and Gotay (1982) in that persistence on the speech task was greater in the expansive relative to the slumped posture. In addition, Nair et al. discovered that expansive postures may provide protective effects to stress. Upright participants reported feeling more enthusiastic, excited, and strong relative to slumped participants after the stress test. Slumped participants reported feeling more fearful and nervous after the test. Upright participants also reported high self-esteem following the stress test, with less fear of the test. Together, these results highlight the motivational effects of these postures and indicate that expansive postures may help maintain approach motivation even when persons are confronted with stressors.

Other research has investigated manipulated postures and tested the idea that reclining backward might be associated with low approach motivation, as suggested by the fact that reclining often occurs in post-goal states. For example, reclining often occurs after a goal is accomplished or during relaxation, such as after eating a delicious meal. In an initial experiment (E. Harmon-Jones & Peterson, 2009), participants wrote an essay they believed would be evaluated by another participant. Participants were instructed to recline in a chair or sit upright before receiving feedback on the essay. All of the participants in the reclining condition and half in the upright condition received insulting feedback on their essay. The other half of participants in the upright condition received neutral feedback. Results demonstrated that upright participants who were insulted relative to not insulted had greater left frontal cortical activity, a finding that conceptually replicated past experiments (E. Harmon-Jones & Sigelman, 2001). The novel finding indicated that reclining participants who received insulting feedback had less relative left frontal cortical activity than insulted upright participants. These findings suggest that reclining backward may reduce approach-motivated anger. Such findings are consistent with prophesizing by Muhammad, who over 1,400 years ago suggested that anger may be alleviated by sitting down and, if not, then by lying down (Sunan Abu-Dawud, Book 41, No. 4764).

These results were further extended by Price and Harmon-Jones (2010a). Reclining in a chair may be associated with lower approach motivation, whereas leaning forward with arms outstretched in a chair may be associated with higher approach motivation, as when an organism desires to move toward a pleasing object, such as a delicious meal. An upright sitting posture also was included and hypothesized to fall between these other two conditions. EEG was recoded for 1 min while participants maintained these postures. A linear trend effect was found, with reclining producing less relative left frontal activity than leaning forward. The upright posture fell between these two conditions, as predicted.

Thus far, only postures by themselves and in regard to negative insults had been tested. Another experiment tested if whole-body postures influenced left frontal cortical activity to appetitive/desirable stimuli. Recall that demonstrating changes in frontal cortical activity to picture stimuli has produced mixed results. One potential reason for this finding is that pictures likely evoke weak appetitive states. One study demonstrated that positive films influenced relative left frontal activity only when participants formed smiles while watching the films (Davidson et al., 1990). To examine whether whole-body manipulations would have the same effect, E. Harmon-Jones, Gable, and Price (2011) used only the extremes in motivational body postures: leaning and reclining. In each posture, participants viewed neutral rock and appetitive dessert pictures. Results indicated that leaning elicited greater left frontal activity to appetitive dessert pictures as compared to neutral rock pictures. Reclining, however, did not produce a difference in left frontal activity to the two forms of stimuli. These results suggest that motivational postures can influence asymmetric activity to appetitive but not neutral pictures.

The neutral picture results may appear inconsistent with the result that postures can influence left frontal cortical activity while participants are in a resting or baseline-neutral state (Price & Harmon-Jones, 2010a). We suspect subtle effects of posture in a resting state. Even when a neutral stimulus is presented, however, the neutral stimulus may override the subtle effect of resting state postures.

Whole-body postures also influence motivated attention (i.e., LPPs) to emotive stimuli. Price, Dieckman, and Harmon-Jones (2012) again used the extremes in postures: leaning and reclining. In each posture, participants viewed images matched for composition: men and women in erotic situations (appetitive) or walking and talking in public (neutral). Results indicated that leaning postures relative to reclining postures produced larger LPPs to erotica. Postures did not influence LPPs to neutral stimuli, however. In addition, the postures influenced even earlier ERPs associated with rapid processing of visual stimuli, the P1. The P1 is influenced by low-level stimulus features (Pourtois et al., 2005) and increases as stimuli become increasingly visible (Schupp et al., 2008). These results suggest that heightened approach motivation elicited by leaning forward influences how individuals process even the very early stimulus features of motivationally significant images.

Other motivational direction research has focused on leaning behavior. Fawver, Shinichi, Hass, and Janelle (2012) placed participants on a force plate and measured their anterior lean toward affective pictures (i.e., attack, happy, contamination, erotica, and multination stimuli) displayed several feet in front of them. Results indicated that images of attack and aggression led participants to lean forward more than all other stimuli conditions, consistent with the notion that anger is often an approach-oriented emotion.

Finally, research has demonstrated that the motivational direction of emotional states influences how individuals actually move. Fawver, Beatty, Naugle, and Janelle (2014) had participants autobiographically recall different emotional states associated with approach and avoidance motivations. Thereafter, participants walked across the room. The researchers measured initial walking gait, walking speed, and the force of each step. Results indicated that approach-oriented emotional memories (i.e., happiness, anger) relative to avoidant-oriented emotional memories (i.e., fear, sadness) produced more gait initiation, more forceful steps, and faster walking speeds. These results provided robust evidence that forward postural motions are tightly connected with the motivational direction of emotions.

Arm and Hand Postures

Other embodiment research on motivational direction has used partial body postures, such as hand contractions. The motor cortex and the frontal cortex share strong cortical connections (E. Harmon-Jones, 2006; Schiff & Lamon 1989, 1994). Importantly, the primary motor cortex has no corpus callosal connections, which ensures that activation of movement on one side of the body does not produce movement on the other side (Mesulam, 2000). Thus, moving one side of the body, through spreading of activation, may influence lateralized cortical activity within the PFC. However, sensory and motor pathways are crossed (Rinn, 1984). Thus, moving the left side of the body involves activation of the right region of the motor cortex and this activation subsequently may activate regions of the right PFC, with the opposite effect for right-sided body movements.

Initial research tested if unilateral hand contractions might influence motivational response using EEG asymmetry (E. Harmon-Jones, 2006). In this experiment, right-handed participants squeezed a ball with either their left or right hand for two 45s periods of time (with 15s of rest in between) and again while listening to a mildly positive approach-oriented pilot radio broadcast covering living options for the participants. (Right-handed participants were chosen, as left-handed participants may show different patterns of activation.) Results indicated that right-handed contractions compared to left-handed contractions elicited greater left frontal cortical activity. EEG over the motor cortex showed similar results. Importantly, right-handed contractions increased positive approach affect, as measured by the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988), to the radio broadcast. Thus, hand contractions may influence the motivational state of an individual.

These results were extended by Peterson, Shackman, and Harmon-Jones (2008). In this experiment, participants wrote an essay on a controversial topic. Similar to the experiment reviewed earlier, participants then received insulting feedback on their essay from another (ostensible) participant. Prior to receiving feedback, however, participants squeezed a ball with either their left or right hand in order to increase right or left frontal cortical activity. Next, participants played a reaction time game against the "participant" who had insulted them. This was, in fact, a chance to aggress against the participant who had insulted them. Participants delivered noise blasts varying in intensity and duration to the insulting participant. Results revealed that participants who made right-handed contractions gave louder and longer noise blasts compared to participants who made left-handed contractions. In addition, these aggressive responses were correlated with greater relative left frontal activity in the right contraction condition. Finally, results indicated that left and right contractions affected the motor cortex in different ways. Left-handed contractions elicited greater coherence between the motor cortex and posterior regions, whereas right contractions elicited greater coherence between the motor cortex and prefrontal regions. Together, these results suggested that hand contractions can have differential effects on prefrontal asymmetry and approach-oriented emotional states.

Besides influencing approach motivation, hand contractions have been demonstrated to elicit avoidance motivation. Research on stereotype threat—the threat experienced by individuals facing in-group stereotypes that may explain poor performance in high-stakes settings—also has been conducted. Stereotype threat has been demonstrated to be reduced by increasing avoidance motivation. For example, when stereotype-threatened individuals framed a task in terms of losses, which is congruent with the goal of avoiding failure, they showed better performance (Grimm, Markman, Maddox, & Baldwin, 2009). To test the idea that left-hand contractions (in right-handers) would induce avoidance motivation and improve performance in stereotype-threatened individuals, Chalabaev, Radel, Masicampo, and Dru (2016) conducted four experiments. In the first experiment, college students completed a difficult arithmetic task. Participants were informed beforehand that students in their major performed worse on the test relative to other students (stereotype threat) or that they performed better (control condition). While completing the task, right-handed participants squeezed a foam ball with either the left or the right hand. Results indicated that squeezing the ball with the left hand improved performance on the task under stereotype threat. Further experiments from a grounded cognitive perspective demonstrated that presenting math problems on the left versus the right side of a computer screen also reduced stereotype threat (Experiment 2); conceptually similar results occurred with different threats and tasks (Experiment 3 and 4).

Other experiments have used different arm postures. Arm crossing is thought to be indicative of avoidant behaviors (Fetterman, 2015, Experiment 1). Results indicated that covert manipulations to induce arm crossing led to feelings of submissiveness and social vulnerability in participants (Experiment 2). Results further suggested that arm crossing influenced an individual's response to interpersonal violence. That is, results (Experiment 3) indicated that covert arm crossing relative to non-crossed arms at one's sides (control condition) increased self-reported desires to escape a violent interpersonal scenario.

Other embodiment research on motivational direction has used different arm postures. Past research had demonstrated that flexing the arm as though to bring desired stimuli toward the body led to the formation of positive attitudes toward neutral stimuli, but extending the arm as if to push stimuli away from the body led to the formation of negative attitudes toward neutral stimuli (Cacioppo, Priester, & Berntson, 1993).

Deuter, Best, Kuehl, Neumann, and Schachinger (2014) tested to see if these motivational arm postures would further influence the emotive modulated startle reflex. In this experiment, participants viewed appetitive (e.g., erotica) and aversive (e.g., mutilations) pictures. A fixation cross initiated each trial. Following fixation, participants were instructed to flex or extend their arms. Afterward, an affective picture was displayed with a startle stimulus presented during picture viewing. Results indicated a linear trend effect whereby positive stimuli attenuated startle responses more so than neutral and then aversive stimuli. In addition, the act of flexing the arm versus pushing away with an extended arm increased startle responses to negative stimuli. Flexing the arm to pull negative stimuli toward the self is inconsistent with the normal response of pushing negative elements away from the self, which explains why startle responses were magnified in this instance.

These results suggest that hand contractions and arm postures have the ability to influence approach and avoidance motivation, potentially through a mechanism between the motor and the prefrontal cortices. Furthermore, these experiments reinforced the finding that anger often is an approach-oriented emotion.

MOTIVATION TO REGULATE CONFLICTING MOTIVATIONS

Motivations often come into conflict with one another. Many psychological perspectives have considered how conflicting motivations are regulated. One theory that has received a tremendous amount of research for over 70 years is cognitive dissonance theory. This theory is often considered a "cognitive consistency" theory, even though Beauvois and Joule (1996, 1999) have presented convincing arguments against this assertion. Instead, they and others have posited that dissonance theory most often concerns the justification of behavioral commitments. Following this tradition, the action-based model of cognitive dissonance posits that cognitive dissonance processes are primarily aimed at resolving "cognitive conflicts." As cognitions serve to guide and direct action, cognitive conflicts may be thought of more accurately as conflicts between action tendencies. By selectively strengthening particular cognitions while reducing the strength of others, individuals support specific courses of action. Next we briefly describe the conceptual history of cognitive dissonance and then present the action-based model and research derived from it.

Cognitive dissonance theory originally was proposed by Leon Festinger in the 1950s. His theory stated that, when a person holds two cognitions (elements of knowledge) that are both relevant to one another and inconsistent (meaning that the veracity of one suggests that the other should not be true), an uncomfortable psychological state results. Festinger called both the discrepancy between cognitions and the resultant discomfort "dissonance," and he proposed that the magnitude of the dissonance depends on the number and importance of cognitions consonant with, and dissonant with, a particular "generative" cognition (Festinger, 1957). He also proposed that this unpleasant state motivates individuals to engage in psychological work to reduce the discrepancy between cognitions. Commonly, this work involves changing one or more attitudes to reduce the discrepancy. Attitude change is in the direction of the cognition that is most resistant to change, which, in laboratory experiments, is often assumed to be a cognition about recent behavior. This is because, once a person has behaved, it is difficult to undo that knowledge.

Because individuals must act despite competing goals, many contradictory pieces of information, and incomplete knowledge, dissonance processes are common sources of motivation in real-world situations. The flexibility and applicability of the theory account for its longevity. In addition, Aronson (1992) noted that several other theories in social psychology may be "dissonance in disguise," including symbolic self-completion theory (Wicklund & Gollwitzer, 1982), action identification theory (Vallacher & Wegner, 1987), self-evaluation maintenance theory (Tesser, 1988), self-affirmation theory (Steele, 1988), and self-discrepancy theory (Higgins, 1989). Proulx, Inzlicht, and Harmon-Jones (2012) reviewed a number of theories that bear close similarity to dissonance processes. These perspectives regarding how individuals resolve conflicts among perceptions, beliefs, and other cognitions include Piaget's concept of disequilibrium, the state of inconsistency that motivates changes to schemata (Piaget, 1954); system justification (Jost, Banaji, & Nosek, 2004); and the meaning maintenance model (Heine, Proulx, & Vohs, 2006). Thus, cognitive dissonance theory converges with many other motivational perspectives.

In laboratory tests of dissonance theory, a few paradigms commonly have been used to show that individuals change their attitudes and beliefs to support recent behavioral commitments. The free choice paradigm is predicated on the idea that, following a decision between two attractive alternatives, the cognitions that favor the unchosen alternative or disfavor the chosen alternative are dissonant with the decision. Cognitive work should increase the attractiveness of the chosen alternative, decrease the attractiveness of the unchosen alternative, or both. Research using this paradigm has shown that after making a difficult decision between two attractive alternatives, individuals "spread the alternatives" in this way. However, when they make an easy decision between one attractive and one unattractive alternative, they do not do so (Brehm, 1956).

In the induced compliance paradigm, participants are encouraged to behave contrary to an important attitude. In this situation, being "forced" to engage in the behavior (having low choice) functions as an important consonant cognition for the behavior. Typically, when participants believe they had high choice to engage in the counterattitudinal behavior, they change their original attitude in support of the behavior; when they believe they had low choice, they do not change their original attitude (E. Harmon-Jones, Brehm, Greenberg, Simon, & Nelson, 1996). Less commonly used paradigms are effort justification, in which participants elevate the attractiveness of goals for which they have suffered, and belief disconfirmation, in which participants strengthen beliefs for which they have been exposed to disconfirming evidence.

Several revisions to dissonance theory have been proposed since Festinger (1957) formulated it. Most of these revisions have suggested that violations of the self-concept are a necessary component of dissonance. First, Aronson (1968) proposed self-consistency theory, which stated that dissonance occurs when a person's behavior violates his or her self-concept. Because most individuals have a positive self-concept, dissonance usually is evoked by behavior that is irrational, immoral, or incompetent. The person is then motivated to justify the behavior and restore the positive self-concept. Self-affirmation theory (Steele, 1988) proposed that individuals may not have to justify every act, as long as they can maintain an overall self-image that is morally and adaptively adequate. Thus, research under self-affirmation theory showed that, following counterattitudinal behavior, if individuals are given the opportunity to affirm an unrelated but important value, they will not engage in attitude change. Cooper and Fazio (1984), with their "New Look at Dissonance Theory," proposed an even more restrictive set of criteria for dissonance. They suggested that dissonance is produced only when an individual feels personally responsible for creating an aversive consequence for someone else.

However, mounting evidence recently has emerged that violations of the self-concept are not necessary for dissonance to be evoked. First, dissonance reduction has been observed in nonhuman animals and preschool-age children, both of whom presumably lack a well-developed self-concept. Effort justification has been shown in white rats (Lewis, 1964), pigeons (Zentall, 2010), and grade-school children (Benozio & Diesendruck, 2015). Similarly, spreading of alternatives has been demonstrated in capuchin monkeys and preschoolers (Egan, Santos, & Bloom, 2007). Second, adult participants have displayed attitude change after writing a counterattitudinal essay that was then discarded in the trash before anyone could see it. In one example experiment, participants drank a foul-tasting beverage and then were given high or low choice to write that they liked it, after which they immediately discarded the paper on which they wrote the statement. The participants were asked prior to writing to discard the paper, because the experimenter said the paper was not needed and that the "memory" study simply required participants to write the statement. Results revealed that participants given high choice then reported liking the beverage more than those given low choice. In this and similar studies, the counterattitudinal behavior had no chance of causing an aversive consequence (which typically had been operationalized as harming another person), suggesting that aversive consequences are not necessary to produce dissonance or dissonance reduction (E. Harmon-Jones, 2000; E. Harmon-Jones et al., 1996; E. Harmon-Jones, Harmon-Jones, Serra, & Gable, 2011).

If dissonance is not motivated by violations of the self-concept, then why are humans and other organisms motivated to resolve cognitive inconsistencies? The action-based model of dissonance was proposed to answer this question. The action-based model begins from the assumption that cognitions guide and direct action, so when cognitions conflict with one another, effective action is difficult. Dissonance reduction should, by bringing cognitions into closer alignment, make action more effective and unconflicted. The model suggests that, following a decision, the individual should be in an action-oriented, approach-motivated state (Gollwitzer, 1990; Kuhl, 1984). This approach-motivated state enhances implementation of decisions.

Research on the action-based model of dissonance has shown that individuals in an action-oriented state engage in greater spreading of alternatives (E. Harmon-Jones & Harmon-Jones, 2002); that dissonance reduction is approach-motivated, as measured by left frontal cortical activation (E. Harmon-Jones et al., 2008); and that individuals high in BAS (trait approach motivation) engage in more dissonance reduction (C. Harmon-Jones, Schmeichel, Inzlicht, & Harmon-Jones, 2010). Taken together, the large body of research on cognitive dissonance and related phenomena suggests that managing cognitive inconsistencies is an important motivational force in both humans and nonhuman animals. Ultimately, this motivational force to resolve cognitive inconsistency may be aimed at resolving actional or motivational inconsistencies.

SUMMARY AND CONCLUSION

In this chapter, we reviewed theory and research on two major dimensions of motivation: motivational intensity and motivational direction. In doing so, we reviewed how motivational intensity and direction are related to psychophysiological and cognitive responses. We also reviewed theory and research that suggests that cognitive dissonance processes ultimately may be aimed at resolving motivational conflicts.

Taken together, this review illustrates how the psychophysiological study of motivation is flourishing and increasing the understanding of the relationships among motivation, emotion, and cognition. Although much is now known about motivation and its relationships with other fundamental psychological processes, much remains to be better understood, such as how motivation is caused by influences other than external stimuli. We hope that this review assists in future theoretical and research endeavors.

REFERENCES

Ach, N. (1935). *Analyse des Willens* [Analysis of volitions]. Berlin, Germany: Urban & Schwarzenberg.

Ades, C., Ramires, E. N. (2002). Asymmetry of leg use during prey handling in the spider *Scytodes globula* (Scytodidae). *Journal of Insect Behavior*, *15*, 563–570.

Alema, G., Rosadini, G., & Rossi, G. F. (1961). Preliminary experiments on the effects of the intracarotid introduction of sodium amytal in Parkinsonian syndromes. *Bollettino della Societa Italiana di Biologia Sperimentale*, *37*, 1036–1037.

Aluja, A., Blanch, A., Blanco, E., & Balada, F. (2014). Affective modulation of the startle reflex and the reinforcement sensitivity theory of personality: the role of sensitivity to reward. *Physiology & Behavior*, *138*, 332–339.

Amodio, D. M., & Harmon-Jones, E. (2011). Trait emotions and startle eyeblink responses to affective pictures: on the unique relationship of trait anger. *Emotion*, *11*, 47–51.

Amodio, D. M., Master, S. L., Yee, C. M., & Taylor, S. E. (2008). Neurocognitive components of behavioral inhibition and activation systems: implications for theories of self-regulation. *Psychophysiology*, *45*, 11–19.

Amodio, D. M., Shah, J. Y., Sigelman, J., Brazy, P. C., & Harmon-Jones, E. (2004). Implicit regulatory focus associated with resting frontal cortical asymmetry. *Journal of Experimental Social Psychology*, *40*, 225–232.

Angus, D. J., Kemkes, K., Schutter, D. J., & Harmon-Jones, E. (2015). Anger is associated with reward-related electrocortical activity: evidence from the reward positivity. *Psychophysiology*, *52*, 1271–1280.

Aronson, E. (1968). Dissonance theory: progress and problems. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J.

Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of Cognitive Consistency: A Sourcebook* (pp. 5–27). Chicago, IL: Rand McNally.

Aronson, E. (1992). The return of the repressed: dissonance theory makes a comeback. *Psychological Inquiry*, *3*, 303–311.

Beauvois, J. L., & Joule, R. V. (1996). *A Radical Dissonance Theory*. London, UK: Taylor and Francis.

Beauvois, J. L., & Joule, R. V. (1996). A radical point of view on dissonance theory. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology* (pp. 43–70). Washington, DC: American Psychological Association.

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience*, *34*, 3005–3012. doi: 10.1523/JNEURO SCI.3684–13.2014

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not my fault: event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex*, *20*, 2874–2883. doi:10.1093/cercor/bhq038

Berkman, E. T., & Lieberman, M. D. (2010). Approaching the good and avoiding the bad: separating action and valence using dorsolateral prefrontal cortical asymmetry. *Journal of Cognitive Neuroscience*, *22*, 1970–1979.

Berkowitz, L. (1993). *Aggression: Its Causes, Consequences, and Control*. New York, NY: McGraw Hill.

Black, E. W. (1975). Unilateral brain lesions and MMPI performance: a preliminary study. *Perceptual and Motor Skills*, *40*, 87–93.

Blanchard, R. J., & Blanchard, D. C. (1984). Affect and aggression: an animal model applied to human behavior. In R. J. Blanchard & D. C. Blanchard (Eds.), *Advances in the Study of Aggression* (pp. 1–62). Orlando, FL: Academic Press.

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & van Boxtel, A. (2005). Committee report: guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*, 1–15.

Bradley, M. M. (2009). Natural selective attention: orienting emotion. *Psychophysiology*, *46*, 1–11.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation 1: defensive and appetitive reactions in picture processing. *Emotion*, *1*, 276–298.

Bradley, M. M., & Lang, P. J. (2007). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (3rd ed., pp. 581–607). New York, NY: Cambridge University Press.

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology*, *52*, 384–389.

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Reviews of Psychology*, *40*, 109–131.

Brehm, J. W., Wright, R. A., Solomon, S., Silka, L., & Greenberg, J. (1983). Perceived difficulty, energization, and the magnitude of goal valence. *Journal of Experimental Social Psychology*, *19*, 21–48.

Bress, J. N., & Hajcak, G. (2013). Self-report and behavioral measures of reward sensitivity predict the feedback negativity. *Psychophysiology*, *50*, 610–616. doi:10.1111/psyp.12053

Benozio, A., & Diesendruck, G. (2015). From effort to value preschool children's alternative to effort justification. *Psychological Science*, *26*(9), 1423–1429.

Briggs, K. B., & Martin, F. H. (2009). Affective picture processing and motivational relevance: arousal and valence effects on ERPs in an oddball task. *International Journal of Psychophysiology*, *72*, 299–306.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*, 452–459.

Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality & Social Psychology*, *65*, 5–17.

Carlson, J. M., Foti, D., Harmon-Jones, E., & Proudfit, G. H. (2015). Midbrain volume predicts fMRI and ERP measures of reward reactivity. *Brain Structure and Function*, *220*(3), 1861–1866.

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *NeuroImage*, *57*, 1608–1616. doi: 10.1016/j.neuroimage.2011.05.037

Carver, C. S. (2004). Negative affects deriving from the behavioral approach system. *Emotion*, *4*, 3–22.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: evidence and implications. *Psychological Bulletin*, *135*, 183–204.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chalabaev, A., Radel, R., Masicampo, E. J. C., & Dru, V. (2016). Reducing stereotype threat with embodied triggers: a case of sensorimotor-mental congruence. *Personality and Social Psychology Bulletin*, *42*, 1063–1076.

Coan, J. A., Allen, J. J. B., & Harmon-Jones, E. (2001). Voluntary facial expression and hemispheric asymmetry over the frontal cortex. *Psychophysiology*, *38*, 912–925.

Cook, I. A., O'Hara, R., Uijtdehaage, S. H. J., Mandelkern, M., & Leuchter, A. F. (1998). Assessing the accuracy of topographic EEG mapping for determining local brain function. *Electroencephalography and Clinical Neurophysiology*, *107*, 408–414.

Cooper, A. J., Duke, É., Pickering, A. D., & Smillie, L. D. (2014). Individual differences in reward prediction error: contrasting relations between feedback-related negativity and trait measures of reward sensitivity, impulsivity and extraversion. *Frontiers in Human Neuroscience*, *8*, 248. doi: 10.3389/fnhum.2014.00248

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 17, pp. 229–264). Orlando, FL: Academic Press.

d'Alfonso, A. A. L., van Honk, J., Hermans, E., Postma, A., & de Haan, E. H. F. (2000). Laterality effects in selective attention to threat after repetitive transcranial magnetic stimulation at the prefrontal cortex in female subjects. *Neuroscience Letters*, *280*, 195–198.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London, UK: Murray.

Davidson, R. J., Chapman, J. P., Chapman, L. J., & Henriques, J. B. (1990). Asymmetric brain electrical-activity discriminates between psychometrically-matched verbal and spatial cognitive tasks. *Psychophysiology*, *27*, 528–543.

Davidson, R. J., & Fox, N. (1982). Asymmetric brain activity discriminates between positive and negative affective stimuli in 10 month old infants. *Science*, *218*, 1235–1237.

Davis, M. (2006). Neural systems involved in fear and anxiety measured with fear-potentiated startle. *American Psychologist*, *61*, 741–756.

Deckel, A. W., Lillaney, R., Ronan, P. J., & Summers, C. H., (1998). Lateralized effects of ethanol on aggression and serotonergic systems in *Anolis carolinensis*. *Brain Research*, *807*, 38–46.

Deuter, C. E., Best, D., Kuehl, L. K., Neumann, R., & Schachinger, H. (2014). Effects of approach-avoidance related motor behavior on the startle response during emotional picture processing. *Biological Psychology*, *103*, 292–296.

Duclos, S. E., Laird J. D., Schneider, E., Sexter, M., Stern, L., & Van Lighten, O. (1989). Emotion-specific effects of facial expressions and postures on emotional experience. *Journal of Personality and Social Psychology*, *57*, 100–108.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183–201.

Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance evidence from children and monkeys. *Psychological Science*, *18*(11), 978–983.

Ekman, P., & Davidson, R. J. (1993). Voluntary smiling changes regional brain activity. *Psychological Science*, *4*, 342–345.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Englewood Cliffs, NJ: Prentice-Hall.

Elliott, R. (1969). Tonic heart rate: experiments on the effects of collative variables lead to a hypothesis about its motivational significance. *Journal of Personality and Social Psychology*, *12*, 211–228.

Fawver, B., Beatty, G. F., Naugle, K. M., & Janelle, C. M. (2014). Emotional state impacts center of pressure displacement before forward gait initiation. *Journal of Applied Biomechanics*, *31*, 35–40.

Fawver, B., Shinichi, A., Hass, C. J., & Janelle, C. M. (2012). Active control of approach-oriented posture is influenced by emotional reactions. *Emotion*, *12*, 1350–1361.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Fetterman, A. K., Blair, J. L., & Robinson, M. D. (2015). Submissive, inhibited, avoidant, and escape motivated: the correlates and consequences of arm-crossing. *Motivation Science*, *1*, 37–46.

Ford, B. Q., Tamir, M., Brunyé, T. T., Shirer, W. R., Mahoney, C. R., & Taylor, H. A. (2010). Keeping your eyes on the prize: anger and visual attention to threats and rewards. *Psychological Science*, *21*, 1098–1105.

Ford, B. Q., Tamir, M., Gagnon, S. A., Taylor, H. A., & Brunyé, T. T. (2012). The angry spotlight: trait anger and selective visual attention to rewards. *European Journal of Personality*, *26*, 90–98.

Foti, D., & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: evidence from event-related potentials. *Biological Psychology*, *81*, 1–8. doi: 10.1016/j.biopsycho.2008.12.004

Foti, D., & Hajcak, G. (2010). State sadness reduces neural sensitivity to nonrewards versus rewards. *NeuroReport*, *21*, 143–147. doi: 10.1097/WNR.0b013e3283356448

Foti, D., Kotov, R., Klein, D. N., & Hajcak, G. (2011). Abnormal neural sensitivity to monetary gains versus losses among adolescents at risk for depression. *Journal of Abnormal Child Psychology*, *39*, 913–924. doi: 10.1007/s10802-011-9503-9

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: the broaden and build theory of positive emotions. *American Psychologist*, *56*, 218–226.

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion*, *19*, 313–332.

Frijda, N. H. (1986). *The Emotions*. New York, NY: Cambridge University Press.

Gable, P. A., & Harmon-Jones, E. (2008). Approach motivated positive affect reduces breadth of attention. *Psychological Science*, *19*, 476–482.

Gable, P. A., & Harmon-Jones, E. (2010a). The blues broaden, but the nasty narrows: attentional consequences of negative affects low and high in motivational intensity. *Psychological Science*, *21*, 211–215.

Gable, P. A., & Harmon-Jones, E. (2010b). The effect of low vs. high approach-motivated positive affect on memory for peripherally vs. centrally presented information. *Emotion*, *4*, 599–603.

Gable, P. A., & Harmon-Jones, E. (2010c). Late positive potential to appetitive stimuli and local attentional bias. *Emotion*, *10*, 441–446.

Gable, P. A., & Harmon-Jones, E. (2010d). The motivational dimensional model of affect: implications for breadth of attention, memory, and cognitive categorization. *Cognition and Emotion*, *24*, 322–337.

Gable, P. A., & Harmon-Jones, E. (2013). Does arousal per se account for the influence of appetitive stimuli on attentional scope and the late positive potential? *Psychophysiology*, *50*, 344–350.

Gable, P. A., & Poole, B. D. (2014). Influence of trait behavioral inhibition and behavioral approach motivation systems on the LPP and frontal asymmetry to anger pictures.

Social Cognitive and Affective Neuroscience, *9*, 182–190.

Gable, P. A., Poole, B. D., & Harmon-Jones, E. (2015). Anger narrows perceptual and conceptual cognitive scope. *Journal of Personality and Social Psychology*, *109*, 163–174.

Gainotti, G. (1972). Emotional behavior and hemispheric side of the lesion. *Cortex 8*, 41–55.

Gard, D. E., Gard, M. G., Mehta, N., Kring, A. M., & Patrick, C. J. (2007). Impact of motivational salience on affect modulated startle at early and late probe times. *International Journal of Psychophysiology*, *66*, 266–270.

Gasparrini, W. G., Satz, P., Heilman, K., & Coolidge, F. L. (1978). Hemispheric asymmetries of affective processing as determined by the Minnesota Multiphasic Personality Inventory. *Journal of Neurology, Neurosurgery and Psychiatry*, *41*, 470–473.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282. doi:10.1126/science.1066893

Gendolla, G. H., & Wright, R. A. (2012). Effort intensity: some insights from the cardiovascular system. In R. M. Richard (Ed.), *The Oxford Handbook on Motivation* (pp. 420–438). New York, NY: Oxford University Press.

Glass, D. C., & Singer, D. C. (1972). *Urban Stress: Experiments on Noise and Social Stressors*. New York, NY: Academic Press.

Goldstein, K. (1939). *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man*. Boston, MA: Beacon Press.

Gollwitzer, P. M. (1990). Action phases and mindsets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (Vol. 2, pp. 53–92). New York, NY: Guilford Press.

Gomez, A., & Gomez, R. (2002). Personality traits of the behavioral approach and inhibition systems: associations with processing of emotional stimuli. *Personality and Individual Differences*, *32*, 1299–1316.

Grimm, L. R., Markman, A. B., Maddox, W. T., & Baldwin, G. C. (2009). Stereotype threat reinterpreted as a regulatory mismatch. *Journal of Personality and Social Psychology*, *96*, 288–304. doi:10.1037/a0013463

Gu, R., Ge, Y., Jiang, Y., & Luo, Y.-J. (2010). Anxiety and outcome evaluation: the good, the bad and the ambiguous. *Biological Psychology*, *85*, 200–206. doi: http://dx.doi.org/10.1016/j.biopsycho.2010.07.001

Gu, R., Huang, Y. X., & Luo, Y. J. (2010). Anxiety and feedback negativity. *Psychophysiology*, *47*, 961–967. doi: 10.1111/j.1469–8986.2010.00997.x

Güntürkün, O., Diekamp, B., Manns, M., Nottelmann, F., Prior, H., Schwarz, A., & Skiba, M. (2000). Asymmetry pays: visual lateralization improves discrimination success in pigeons. *Current Biology*, *10*, 1079–1081.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*, 148–154. doi: 10.1016/j.biopsycho.2005.04.001

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2011). ERPs and the study of emotion. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford Handbook of ERP Components* (pp. 441–474). New York, NY: Oxford University Press.

Harmon-Jones, C., Schmeichel, B. J., Inzlicht, M., & Harmon-Jones, E. (2011). Trait approach motivation relates to dissonance reduction. *Social Psychological and Personality Science*, *2*, 21–28.

Harmon-Jones, C., Schmeichel, B. J., Mennitt, E., & Harmon-Jones, E. (2011). The expression of determination: similarities between anger and approach-related positive affect. *Journal of Personality and Social Psychology*, *100*, 172–181.

Harmon-Jones, E. (2000). Cognitive dissonance and experienced negative affect: evidence that dissonance increases experienced negative affect even in the absence of aversive consequences. *Personality and Social Psychology Bulletin*, *26*, 1490–1501.

Harmon-Jones, E. (2003). Anger and the behavioral approach system. *Personality and Individual Differences*, *35*, 995–1005. http://dx.doi.org/10.1016/S0191-8869(02)00313-6

Harmon-Jones, E. (2004). On the relationship of anterior brain activity and anger: examining the role of attitude toward anger. *Cognition and Emotion*, *18*, 337–361.

Harmon-Jones, E. (2006). Unilateral right-hand contractions cause contralateral alpha power suppression and approach motivational affective experience. *Psychophysiology*, *43*, 598–60.

Harmon-Jones, E., & Allen, J. J. B. (1997). Behavioral activation sensitivity and resting frontal EEG asymmetry: covariation of putative indicators related to risk for mood disorders. *Journal of Abnormal Psychology*, *106*, 159–163.

Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and prefrontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology*, *74*, 1310–1316.

Harmon-Jones, E., Brehm, J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Evidence that the production of aversive consequences is not necessary to create cognitive dissonance. *Journal of Personality and Social Psychology*, *70*, 5–16.

Harmon-Jones, E., & Gable, P. A. (2009). Neural activity underlying the effect of approach-motivated positive affect on narrowed attention. *Psychological Science*, *20*, 406–409.

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2011). Leaning embodies desire: evidence that leaning forward increases relative left frontal cortical activation to appetitive stimuli. *Biological Psychology*, *87*, 311–313.

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2013). Does negative affect always narrow and positive affect always broaden the mind? Considering the influence of motivational intensity on cognitive scope. *Current Directions in Psychological Science*, *22*, 301–307.

Harmon-Jones, E., & Harmon-Jones, C. (2002). Testing the action-based model of cognitive dissonance: the effect of action-orientation on post-decisional attitudes. *Personality and Social Psychology Bulletin*, *28*, 711–723.

Harmon-Jones, E., Harmon-Jones, C., Fearn, M., Sigelman, J. D., & Johnson, P. (2008). Action orientation, relative left frontal cortical activation, and spreading of alternatives: a test of the action-based model of dissonance. *Journal of Personality and Social Psychology*, *94*, 1–15.

Harmon-Jones, E., Harmon-Jones, C., & Price, T. F. (2013). What is approach motivation? *Emotion Review*, *5*, 291–295.

Harmon-Jones, E., Harmon-Jones, C., Serra, R., & Gable, P. A. (2011). The effect of commitment on relative left frontal cortical activity: tests of the action-based model of dissonance. *Personality and Social Psychology Bulletin*, *37*, 395–408.

Harmon-Jones, E., & Peterson, C. K. (2009). Supine body position reduces neural response to anger evocation. *Psychological Science*, *20*, 1209–1210.

Harmon-Jones, E., Peterson, C. K., & Harris, C. R. (2009). Jealousy: novel methods and neural correlates. *Emotion*, *9*, 113–117.

Harmon-Jones, C., Schmeichel, B. J., Mennitt, E., & Harmon-Jones, E. (2011). The expression of determination: similarities between anger and approach-related positive affect. *Journal of Personality and Social Psychology*, *100*, 172–181.

Harmon-Jones, E., & Sigelman, J. D. (2001). State anger and prefrontal brain activity: evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology*, *80*, 797–803.

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. doi: http://dx.doi.org/10.1016/j.neuroimage.2013.08.028

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Research*, *1105*, 93–101. doi: 10.1016/j.brainres.2005.12.015

Hawk, L. W., & Kowmas, A. D. (2003). Affective modulation and prepulse inhibition of startle among undergraduates high and low in behavioral inhibition and approach. *Psychophysiology*, *40*, 131–138.

Heerey, E. A. (2014). Learning from social rewards predicts individual differences in self-reported social ability. *Journal of Experimental Psychology: General*, *143*, 332–339.

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: on the coherence of social motivations. *Personality and Social Psychological Review*, *10*, 88–111.

Henriques, J. B., & Davidson, R. J. (1990). Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects. *Journal of Abnormal Psychology*, *99*, 22–31.

Hewig, J., Hagemann, D., Seifert, J., Naumann, E., & Bartussek, D. (2004). On the selective relation of frontal cortical activity and anger-out versus anger-control. *Journal of Personality and Social Psychology*, *87*, 926–939.

Hicks, J., Friedman, R. S., Gable, P. A., & Davis, W. E. (2012). Interactive effects of approach motivational intensity and alcohol cues on the scope of perceptual attention. *Addiction*, *107*, 1074–1080.

Higgins, E. T. (1989). Self-discrepancy theory: what patterns of self-beliefs cause people to suffer? In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 22, pp. 93–136). San Diego, CA: Academic Press.

Hillgruber, A. (1912). *Fortlaufende Arbeit und Willensbetatigung* [Continuous work and will performance]. *Untersuchungen zür Psychologie und Philosophie* [Studies on the Psychology and Philosophy], *1*.

Hopkins, W. D., Bennett, A. J., Bales, S.L., Lee, J., & Ward, J. P. (1993). Behavioral laterality in captive bonobos (*Pan paniscus*). *Journal of Comparative Psychology*, *107*, 403–410.

Isen, A. M., & Daubman, K. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology*, *47*, 1206–1217.

Izard, C. E. (1991). *The Psychology of Emotions*. New York, NY: Plenum Press.

Jacobs, G. D., & Snyder, D. (1996). Frontal brain asymmetry predicts affective style in men. *Behavioral Neuroscience*, *110*, 3–6.

James, W. (1890). *The Principles of Psychology* (Vol. 2). New York, NY: Dover Publications.

Jensen-Campbell, L. A., Knack, J. M., Waldrip, A. M., & Campbell, S. D. (2007). Do big five personality traits associated with self-control influence the regulation of anger and aggression? *Journal of Research in Personality*, *41*, 403–424.

Jones, N. A., & Fox, N. A. (1992). Electroencephalogram asymmetry during emotionally evocative films and its relation to positive and negative affectivity. *Brain and Cognition*, *20*, 280–299.

Jost, J., Banaji, M., & Nosek, B. (2004). A decade of system justification theory: accumulated evidence of conscious and unconscious bolstering of the status. *Political Psychology*, *25*, 881–920.

Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large scale neural correlates of affective picture processing. *Psychophysiology*, *39*, 641–649.

Kimchi, R., & Palmer, S. E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 521–535.

Knutson, B., & Wimmer, G. E. (2007). Reward: neural circuitry for social evaluation. In E. Harmon-Jones & P. Winkielman (Eds.), *Social Neuro-Science: Integrated Biological and Psychological Explanations of Social Behavior* (pp. 157–175). New York, NY: Guilford Press.

Kuhl, J. (1984). Volitional aspects of achievement motivation and learned helplessness: toward a comprehensive theory of action-control. In B. A. Maher (Ed.), *Progress in Experimental Personality Research* (Vol. 13, pp. 99–171), New York, NY: Academic Press.

Kujawa, A., Smith, E., Luhmann, C., & Hajcak, G. (2013). The feedback negativity reflects favorable compared to nonfavorable outcomes based on global, not local, alternatives. *Psychophysiology*, *50*, 134–138. doi: 10.1111/psyp.12002

Lagerspetz, K. M. J. (1969). Aggression and aggressiveness in laboratory mice. In S. Garattini & E. B. Sigg (Eds.), *Aggressive Behavior* (pp. 77–85). New York, NY: Wiley.

Laird, J. D. (1974). Self-attribution of emotion: the effects of expressive behavior on the quality of emotional experience. *Journal of Personality and Social Psychology*, *29*, 475–486.

Lang, P. J. (1985). The cognitive psychophysiology of emotion: fear and anxiety. In A. H. Tuma, and J. D. Maser (Eds.), *Anxiety and the Anxiety Disorders* (pp. 131–170). Hillsdale, NJ: Erlbaum.

Lang, P. J. (1995). The emotion probe. *American Psychologist*, *50*, 372–385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*, 377–398.

Lewis, M. (1964). Some nondecremental effects of effort. *Journal of Comparative and Physiological Psychology*, *57*(3), 367.

Lewis, M., Alessandri, S. M., & Sullivan, M. W., (1990). Violation of expectancy, loss of control, and anger expressions in young infants. *Developmental Psychology*, *26*, 745–751.

Lewis, M., Sullivan, M. W., Ramsey, D. S., & Alessandri, S. M. (1992). Individual differences in anger and sad expressions during extinction: antecedents and consequences. *Infant Behavior and Development*, *15*, 443–452.

MacLeod, C. (1999). Anxiety and anxiety disorders. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 447–477). New York, NY: Wiley.

Maltzman, I. (1979). Orienting reflexes and classical conditioning in humans. In H. D. Kimmel, E. H. van Olst, & J. F. Orlebeke (Eds.), *The Orienting Reflex in Humans* (pp. 323–351). New York, NY: Erlbaum.

Matheson, H. E., & Barsalou, L. (2017). Embodied cognition. In J. T. Wixted & S. Thompson-Schill (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 3: *Language and Thought*). New York, NY: Wiley.

Mesulam, M. M. (2002). *Principles of Behavioral and Cognitive Neurology*. New York, NY: Oxford University Press.

Mühlberger, C., Angus, D., Jonas, E., Harmon-Jones, C., & Harmon-Jones, E. (2017). Perceived control increases the reward positivity and stimulus preceding negativity. *Psychophysiology*, *54*, 310–322. doi:10.1111/psyp.12786

Nair, S., Sagar, M., Sollers, J., Consedine, N., & Broadbent, E. (2015). Do slumped and upright postures affect stress responses? A randomized trial. *Health Psychology*, *34*, 632–641.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Nobre, K. (2017). Attention. In J. T. Wixted, E. Phelps, & L. Davachi (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 1: *Learning & Memory).* Hoboken, NJ: Wiley.

Nusslock, R., Harmon-Jones, E., Alloy, L. B., Urosevic, S., Goldstein, K., & Abramson, L. Y. (2012). Elevated left mid-frontal cortical activity prospectively predicts conversion to bipolar I disorder. *Journal of Abnormal Psychology*, *121*, 592–601.

Nusslock, R., Shackman, A. J., Harmon-Jones, E., Alloy, L. B., Coan, J. A., & Abramson, L. Y. (2011). Cognitive vulnerability and frontal brain asymmetry: common predictors of first prospective depressive episode. *Journal of Abnormal Psychology*, *120*, 497–503.

Obrist, P. A. (1976). The cardiovascular-behavioral interaction as it appears today. *Psychophysiology*, *13*, 95–107.

Papillo, J. F., & Shapiro, D. (1990). The cardiovascular system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of Psychophysiology* (pp. 465–512). New York, NY: Cambridge University Press.

Perria, P., Rosadini, G., & Rossi, G. F. (1961). Determination of side of cerebral dominance with amobarbital. *Archives of Neurology*, *4*, 175–181.

Peterson, C. K., Shackman, A. J., & Harmon-Jones, E. (2008). The role of asymmetrical frontal cortical activity in aggression. *Psychophysiology*, *45*, 86–92.

Piaget, J. (1954). *The construction of reality in the child*. New York, NY: Basic Books. (Originally published 1937.)

Pizzagali, D. A., Sherwood, R. J., Henriques, J. B., & Davidson, R. J. (2005). Frontal brain asymmetry and reward responsiveness. *Psychological Science*, *16*, 805–813.

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. New York, NY: Harper Collins.

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. *Human Brain Mapping*, *26*, 65–79.

Prause, N., Steele, V. R., Staley, C., & Sabatinelli, D. (2014). Late positive potential to explicit sexual images associated with the number of sexual intercourse partners. *Social Cognitive and Affective Neuroscience*, *10*, 93–100.

Price, T. F., Dieckman, L., & Harmon-Jones, E. (2012). Embodying approach motivation: body posture influences startle eyeblink and event-related potential responses to appetitive stimuli. *Biological Psychology*, *90*, 211–217.

Price, T. F., & Harmon-Jones, E. (2010a). Approach motivational body postures lean towards left frontal brain activity. *Psychophysiology*, *48*, 718–722.

Price, T. F., & Harmon-Jones, E. (2010b). The effect of embodied emotive states on cognitive categorization. *Emotion*, *10*, 934–938.

Price, T. F., Hortensius, R., & Harmon-Jones, E. (2013). Neural and behavioral associations of manipulated determination facial expressions. *Biological Psychology*, *94*, 221–227.

Proudfit, G. H. (2015). The reward positivity: from basic research on reward to a biomarker for depression. *Psychophysiology*, *52*, 449–459. doi:10.1111/psyp.12370

Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences*, *16*, 285–291. doi:10.1016/j.tics.2012.04.002

Reeve, J. (2016). A grand theory of motivation: Why not? *Motivation and Emotion*, *40*, 31–35.

Richter, M., Friedrich, A., & Gendolla, G. H. E. (2008). Task difficulty effects on cardiac activity. *Psychophysiology*, *45*, 869–875.

Rinn, W. E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expression. *Psychological Bulletin*, *95*, 52–77.

Riskind, J. H., & Gotay, C. C. (1982). Physical posture: could it have regulatory or feedback effects on motivation and emotion? *Motivation & Emotion*, *6*, 273–298.

Robinson, R. G., & Price, T. R. (1982). Post-stroke depressive disorders: a follow-up study of 103 patients. *Stroke*, *13*, 635–641.

Rogers, L. J. (2002). Lateralised brain function in anurans: comparison to lateralization in other vertebrates. *Laterality: Asymmetries of Body, Brain and Cognition*, *7*, 219–239.

Rossi, G. F., & Rosadini, G. R. (1967). Experimental analyses of cerebral dominance in man. In D. H. Millikan & F. L. Darley (Eds.), *Brain Mechanisms Underlying Speech and Language* (pp. 167–184). New York, NY: Grune & Stratton.

Rybak, M., Crayton, J. W., Young, I. J., Herba, E., & Konopka, L. M. (2006). Frontal alpha power asymmetry in aggressive children and adolescents with mood and disruptive behavior disorders. *Clinical EEG and Neuroscience*, *37*, 16–24.

Sabatinelli, D., Lang, P. J., Keil, A., & Bradley, M. M. (2007). Emotional perception: correlation of functional MRI and event-related potentials. *Cerebral Cortex*, *17*, 1085–1091.

Sackeim, H., Greenberg, M. S., Weimen, A. L., Gur, R. C., Hungerbuhler, J. P., & Geschwind, N., (1982). Hemispheric asymmetry in the expression of positive and negative emotions: neurologic evidence. *Archives of Neurology*, *39*, 210–218.

Santesso, D. L., Bogdan, R., Birk, J. L., Goetz, E. L., Holmes, A. J., & Pizzagalli, D. A. (2012). Neural responses to negative feedback are related to negative emotionality in healthy adults. *Social Cognitive and Affective Neuroscience*, *7*, 794–803. doi:10.1093/scan/nsr054

Schaffer, C. E., Davidson, R. J., & Saron, C. (1983). Frontal and parietal electroencephalogram asymmetry in depressed and nondepressed subjects. *Biological Psychiatry*, *18*, 753–762.

Schiff, B. B., & Lamon, M. (1989). Inducing emotion by unilateral contraction of facial muscles: a new look at hemispheric specialization and the experience of emotion. *Neuropsychologia*, *27*, 923–935.

Schiff, B. B., & Lamon, M. (1994). Inducing emotion by unilateral contraction of hand muscles. *Cortex*, *30*, 247–254.

Schöne, B., Schomberg, J., Gruber, T., & Quirin, M. (2016). Event-related frontal alpha asymmetries: electrophysiological correlates of approach motivation. *Experimental Brain Research*, *234*, 559–567.

Schupp, H. [T.], Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., & Lang, P. (2004). Brain processes in emotional perception: motivated attention. *Cognition & Emotion*, *18*, 593–611.

Schupp, H. T., Stockburger, J., Schmalzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: brain potentials and stimulus identification. *Neuroreport*, *19*, 16–171.

Schutter, D. J. L. G. (2009). Transcranial magnetic stimulation. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in Social Neuroscience* (pp. 233–258). New York, NY: Guilford Press.

Schutter, D. J. L. G., van Honk, J., d'Alfonso, A. A. L., Postma, A., & de Haan, E. H. F. (2001). Effects of slow rTMS at the right dorsolateral prefrontal cortex on EEG asymmetry and mood. *Neuroreport*, *12*, 445–447.

Smillie, L. D., Cooper, A. J., & Pickering, A. D. (2011). Individual differences in reward–prediction–error: extraversion and feedback-related negativity. *Social Cognitive and Affective Neuroscience*, *6*, 646–652. doi:10.1093/scan/nsq078

Smith, T. W., Baldwin, M., & Christensen, A. J. (1990). Inter-personal influence as active coping: effects of task difficulty on cardiovascular reactivity. *Psychophysiology*, *27*, 429–437.

Smits, D. J. M., & Kuppens, P. (2005). The relations between anger, coping with anger, and aggression, and the BIS/BAS system. *Personality and Individual Differences*, *39*, 783–793.

Steele, C. M. (1988). The psychology of self-affirmation: sustaining the integrity of the self. In L. Berkowitz (Ed.) *Advances in Experimental Social Psychology* (Vol. 21, pp. 261–302). San Diego, CA: Academic Press.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*, 768–776.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: a biological substrate of the behavioral approach and inhibition systems. *Psychological Science*, *8*, 204–210.

Tamir, M., & Robinson, M. D. (2007). The happy spotlight: positive mood and selective attention to rewarding information. *Personality and Social Psychology Bulletin*, *33*, 1124–1136.

Terzian, H., & Cecotto, C. (1959). Determination and study of hemisphere dominance by means of intracarotid sodium amytal injection in man: II. Electroencephalographic effects. *Bolletino della Societa Italiana Sperimentale*, *35*, 1626–1630.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. *Advances in Experimental Social Psychology*, *21*, 181–227.

Thorndike, E. L. (1911). *Animal Intelligence*. New York, NY: Macmillan.

Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Doss, R. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *Journal of Personality and Social Psychology*, *62*, 676–687.

Tracey, J. L., & Robins, R. W. (2007). The prototypical pride expression: development of a nonverbal behavior coding system. *Emotion*, *4*, 789–801.

Tullet, A. M., Harmon-Jones, E., & Inzlitch, M. (2012). Right frontal cortical asymmetry predicts empathic reactions: support for a link between withdrawal motivation and empathy. *Psychophysiology*, *49*, 1145–1153.

Tullet, A. M., Wildschut, T., Sedikides, C., & Inzlicht, M. (2015). Right-frontal cortical asymmetry predicts increased proneness to nostalgia. *Psychophysiology*, *52*, 990–996.

Uusberg, H., Allik, J., & Hietanen, J. K. (2015). Eye contact reveals a relationship between neuroticism and anterior EEG asymmetry. *Neuropsychologia*, *73*, 161–168.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*, 3–15.

van Honk, J., & Schutter, D. J. L. G. (2006). From affective valence to motivational direction: the frontal asymmetry of emotion revisited. *Psychological Science*, *17*, 963–965.

Verona, E., Sadeh, N., & Curtin, J. J. (2009). Stress-induced asymmetric frontal brain activity and aggression risk. *Journal of Abnormal Psychology*, *118*, 131–145.

von Hecker, U., & Meiser, T. (2005). Defocused attention in depressed mood: evidence from source monitoring. *Emotion*, *5*, 456–463.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: a new measure of emotion? *Journal of Abnormal Psychology*, *97*, 487–491.

Watson, D. (2000). *Mood and Temperament*. New York, NY: Guilford Press.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic Self-Completion*. Hillsdale, NJ: Erlbaum.

Williams, L. A., & DeSteno, D. (2008). Pride and perseverance: the motivational role of pride. *Journal of Personality and Social Psychology*, *94*, 1007–1017.

Wright, R. A. (1984). Motivation, anxiety, and the difficulty of control. *Journal of Personality and Social Psychology*, *46*, 1376–1388.

Wright, R. A. (1996). Brehm's theory of motivation as a model of effort and cardiovascular response. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The Psychology of Action: Linking Cognition and Motivation to Behavior* (pp. 424–453). New York, NY: Guilford Press.

Wright, R. A. (1998). Ability perception and cardiovascular response to behavioral challenge. In M. Kofta, G. Weary, & G. Sedek (Eds.), *Personal Control in Action: Cognitive and Motivational Mechanisms* (pp. 197–232). New York, NY: Plenum Press.

Young, P. T. (1943). *Emotion in Man and Animal: Its Nature and Relation to Attitude and Motive*. New York, NY: Wiley.

Zajonc, R. B., Pietromonaco, P., & Bargh, J. A. (1982). Independence and interaction of affect and cognition. In M. S. Clark & S. T. Fiske (Eds.), *Affect and Cognition: The 17th Annual Carnegie Symposium* (pp. 211–227). Hillsdale, NJ: Erlbaum.

Zentall, T. R. (2010). Justification of effort by humans and pigeons: cognitive dissonance or contrast? *Current Directions in Psychological Science*, *19*(5), 296–300.

CHAPTER 18

Emotion–Cognition Interactions in Memory and Decision Making

MEGAN E. SPEER AND MAURICIO R. DELGADO

INTRODUCTION

We often think of emotion and cognition as being profoundly different. Emotion involves feelings of pleasure and pain that manifest in subjective, motivational, and physiological changes (Cabanac, 2002). In contrast, cognition is a set of basic processes involving thinking, knowing, and perceiving or even more complex processes, such as imagining and remembering (Miller, 2003). This divergence led researchers to study emotion and cognition in isolation for many decades (Lazarus, 1984; Zajonc, 1984), before realizing the extraordinary extent to which they reciprocally interact to guide behavior. Indeed, there has been a recent surge in research examining the neural and psychological mechanisms underlying emotion and its influence on cognitive processes and vice versa. (For reviews, see Lerner, Li, Valdesolo, & Kassam, 2014; Matheson & Barsalou, 2017; Phelps, Lempert, & Sokol-Hessner, 2014.)

From a neural perspective, regions initially thought to be purely "emotional" (e.g., the amygdala) or purely "cognitive" (e.g., the orbitofrontal cortex) are now known to be critical both for processing emotion and for aiding important cognitive functions, such as memory and decision making, respectively (Cohen, 2005; Damasio, 1994; Squire, 2004). In fact, it is neural mechanisms such as these that afford us the ability to detect and later remember emotionally salient stimuli in our environment (Labar & Cabeza, 2006), which is adaptive for survival (LeDoux, 2014). Emotion also affects higher-level cognitive processing, such as decision making, by swaying our choices in sometimes beneficial and other times detrimental ways. Cognition has a strong influence on emotion as well, particularly for exerting cognitive control over our emotional responses (Ochsner & Gross, 2005; Ochsner, Silvers, & Buhle, 2012). Examining how and why emotion–cognition interactions go awry has helped us identify individual differences in vulnerabilities to mood disorders (e.g., depression), which often involve disturbances in both affective and cognitive domains. These insights provide some examples illustrating that many of our behaviors derived from neural and psychological processes involve dynamic interactions between emotion and cognition.

The overarching goal of this chapter is to discuss the neural and psychological mechanisms underlying emotion–cognition interactions in memory and decision making. We first review how emotion modulates memory processes, beginning at the time of memory encoding and extending to how memory is stored and later retrieved. We then review how emotion modulates decision making in terms of predicting future feelings and influencing choices across various contexts. Finally, we discuss how cognition influences emotion, particularly as a way to regulate our emotional responses. By integrating key findings across studies using behavioral, physiological, and neuroimaging evidence, we illustrate that the way we feel (emotion) and how we think (cognition) are both critical drivers of our behavior.

INFLUENCE OF EMOTION ON COGNITION: MEMORY

Memory Encoding and Attention

Emotionally charged stimuli rapidly grab our attention (Easterbrook, 1959). Several studies have shown that emotions bias attention toward events of emotional significance, which ultimately shapes what reaches our awareness (Anderson, 2005). For example, we detect angry faces faster than neutral faces in visual search tasks (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). This is adaptive because it allows us to selectively respond to the most relevant aspects in our environment while simultaneously inhibiting distractions.

Converging animal models and neuroimaging methods (functional magnetic resonance imaging, positron emission tomography) have pinpointed the amygdala as a region central to the processing of emotion, especially for stimuli of an aversive nature (LeDoux, 2014). The amygdala is thought to play a critical role in enhancing the perceptual awareness of emotionally relevant stimuli, given its reciprocal projections to sensory cortical processing regions (Amaral, Behniea, & Kelly, 2003). For instance, humans engage the amygdala to a greater extent for emotional (e.g., fearful faces) compared to neutral stimuli or even for aversive "oddball" stimuli presented in a neutral context (Strange, Henson, Friston, & Dolan, 2000). Lesion studies examining patients with amygdala damage have been particularly invaluable for understanding the link between the amygdala and perceptual encoding. Early case studies demonstrated that patients with bilateral amygdala damage had difficulty recognizing fear in facial expressions or recognizing multiple emotions within a single facial expression, despite healthy individuals having no difficulty (Adolphs, Tranel, Damasio, & Damasio, 1994; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). Although initially thought to be selective for detecting fear (Öhman & Mineka, 2001), the amygdala responds equally well to the successful encoding of highly arousing positive and negative stimuli (Bonnet et al., 2015; Hamann, Ely, Grafton, & Kilts, 1999). This finding suggests a broader role for the amygdala in filtering for stimuli that are motivationally relevant to the organism's current goals and needs (Cunningham & Brosch, 2012).

Emotional selectivity in the amygdala may occur before conscious awareness (Whalen et al., 1998) and may even be separable from attentional focus (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003). This fact ensures that emotionally salient events are preferentially encoded into memory and available for later retrieval. Evidence of this selectivity comes from studies examining the attentional blink effect. Attentional blink paradigms ask participants to selectively attend to two different target stimuli among a rapid stream of distractor stimuli (Raymond, Shapiro, & Arnell, 1992). When attention is limited, healthy participants have a diminished ability to identify a second target when it is presented soon after the first target stimulus. The only exception is when the second target is emotionally arousing. That is, participants are much more likely to identify an emotional second target (e.g., image of snakes and spiders) than a neutral one (e.g., image of flowers and mushrooms) (Öhman, Flykt, & Esteves, 2001). Yet Anderson and Phelps (2001) showed that individuals with amygdala damage lack enhanced perception for arousing words and thus do not show this attentional blink effect.

Together, these findings illustrate that emotion can influence early perceptual encoding through enhanced selectivity toward arousing stimuli in our environment; neural evidence further supports these findings (Lee, Sakaki, Cheng, Velasco, & Mather, 2013).

Memory Consolidation

Emotion can also have an influence at the next stage of memory formation, namely consolidation—a process whereby information about events is stored in the brain. Consolidation unfolds over time rather than being an instantaneous process. Information slowly assimilates into storage to eventually form stable memory traces (McGaugh, 2000). Given this slow process, memories are fragile and can be modified during consolidation. In this way, important events eliciting greater physiological arousal may be more significant for survival and therefore benefit from stronger consolidation that increases their likelihood of being remembered in the future. Memory traces for less important events are likely to weaken and be later forgotten.

Seminal evidence supporting the idea of consolidation has shown that memory in rodents is impaired if disruptions (e.g., shock or protein synthesis inhibitors) occur during the consolidation window (i.e., shortly after encoding), whereas memory is preserved if disruptions occur after longer delays (Duncan, 1949; Flexner, Flexner, & Stellar, 1965). Similarly, in humans, arousing words are better remembered relative to neutral words after a longer delay (1 hr–1 day) rather than a shorter one (immediate) (Kleinsmith & Kaplan, 1963; LaBar & Phelps, 1998). This is even true when equating attentional resource allocation between emotional and neutral stimuli by presenting stimuli in the periphery (Sharot & Phelps, 2004). Yet patients with amygdala damage have similar forgetting rates for arousing and neutral words regardless of the delay interval, providing further evidence that emotional arousal benefits the consolidation process (LaBar & Phelps, 1998).

Emotional arousal is thought to adaptively influence memory consolidation via the release of stress hormones (epinephrine and glucocorticoids) from the adrenal gland in response to the emotional event itself (McGaugh, 2000). Studies in rodents show that stress hormones activate adrenergic receptors in the amygdala (McGaugh & Roozendaal, 2002), which then have a modulatory role in enhancing consolidation in the hippocampus, a primary neural region for memory processing and storage located adjacent to the amygdala in the medial temporal lobe (MTL). Consistent with animal models, damage to the amygdala in humans or administration of drugs that block stress hormones (e.g., β-adrenergic receptor antagonists) results in diminished memory for emotionally arousing information while preserving memory for neutral information (Cahill, Prins, Weber, & McGaugh, 1994), whereas eliciting stress hormone responses via pain or pharmacological modulation immediately after encoding leads to enhanced memory (Cahill & Alkire, 2003; Cahill, Gorski, & Le, 2003).

Many lines of evidence support the modulatory effect of the amygdala on hippocampal consolidation. Not only do we have greater recall rates for emotional stimuli, both positive and negative, as compared to neutral stimuli, but emotional items that engage the amygdala to a greater extent are more likely to be remembered later. (For review, see Hamann, 2001.) Across individuals, those who elicit the most amygdala activity while encoding emotionally arousing stimuli show the greatest memory boost (Cahill et al., 1996). A significant correlation between amygdala and hippocampal activity during emotional encoding has been observed, although correlations cannot speak to the direction of modulation (Dolcos, LaBar, & Cabeza, 2004). A particularly novel study examining individuals with varying degrees of amygdala and hippocampal damage, however, found inverse relationships between amygdala atrophy and hippocampal activity as well as between hippocampal atrophy and amygdala activity during encoding, highlighting the importance of reciprocal connections between these two regions during memory formation (Richardson, Strange, & Dolan, 2004).

The exact mechanism by which arousing information is consolidated into memory still remains unclear, however. It is hypothesized that memory consolidation takes place through a tag-and-capture process (Ballarini, Moncada, Martinez, Alen, & Viola, 2009; Labar & Cabeza, 2006). Memory traces that were initially irrelevant (and thus weak) may be stored temporarily in case this information gains new relevance in the future. Humans show evidence of this retroactive memory enhancement (Dunsmoor, Murty, Davachi, & Phelps, 2015). Specifically, when neutral stimuli (e.g., tools) were paired with shock, there was a selective enhancement of memory for other previously encoded neutral stimuli from the same category (e.g., tools) that were never paired with shock but not for stimuli from another category (e.g., animals), supporting the notion that weak memory traces can be strengthened retroactively once their greater relevance is realized.

Since memory is fragile and prone to modification during the consolidation process, it has further been suggested that we can update old memories with new information when we reactivate them during retrieval—a process referred to as reconsolidation (Lewis, 1979; Nader & Einarsson, 2010; Nader, Schafe, & Le Doux, 2000). Early evidence came from rodent studies demonstrating that well-consolidated fear memories could be modified via impairment to the amygdala, but only if such impairment occurred once the memories were reopened (Nader et al., 2000). Similarly, Schiller and colleagues (2010) showed that reactivated fear memories in humans could be updated with nonfearful stimuli during the reconsolidation window (i.e., shortly after reactivation), preventing the return of fear at later retrieval (i.e., 24 hours later). Once a memory is consolidated, it does not necessarily remain in a fixed state, suggesting the clinical implications of reconsolidation as a potentially efficacious strategy for updating aversive or traumatic memories, which are central to some psychiatric disorders (e.g., posttraumatic stress disorder, anxiety; Schwabe, Nader, & Pruessner, 2014).

Memory Retrieval

Emotional memory enhancement has been well documented across the literature in studies using a range of emotional stimuli (words, stories, pictures, and film clips; for review, see Hamann, 2001). Emotion does not only boost memory. It can also have a powerful influence on our subjective sense of remembering, especially when remembering real-life personal events termed autobiographical memories. For some memories, we might vividly recall people, places, and events with a high level of detail. For other memories, we might only recognize that we have seen something before without recalling the context, such as recognizing a familiar face but not being able to place where the person was met or how one knows them (Yonelinas, 2002).

One simple way memory researchers study how emotion affects our subjective sense of remembering is by asking participants to make remember (Do you vividly remember this?) versus know judgments (Do you recognize this?). What they have found is that memory tends to be greater for emotional stimuli than neutral stimuli when individuals state they vividly remember seeing the stimuli rather than simply recognizing they saw it (i.e., know) (Dewhurst & Parry, 2000; Kensinger & Corkin, 2003; Ochsner, 2000). This is true even when recall rates are similar for emotional and nonemotional information (Ochsner, 2000; Sharot, Delgado, & Phelps, 2004), suggesting that individuals not only have a general memory boost for emotional versus neutral events but that they reexperience emotional events more vividly.

An extreme example of memory recalled with tremendous detail and clarity is called flashbulb memory (Brown & Kulik, 1977). Flashbulb memories are rare, occurring only for emotional events that are highly significant to the individual. As such, individuals often feel as if they have an almost photographic quality to them—similar to the camera's flashbulb, as implied by the name. For instance, people claim to remember the exact details surrounding surprising and catastrophic events such as the assassination of President John F. Kennedy (Brown & Kulik, 1977), the *Challenger* and *Columbia* space shuttle explosions (Neisser & Harsch, 1992), or the verdict of the O. J. Simpson murder trial (Schmolck, Buffalo, & Squire, 2000).

Interestingly, despite the greater sense of clarity and vividness associated with flashbulb memories, they are not necessarily remembered with better accuracy than ordinary events (Brown & Kulik, 1977; S. R. Schmidt, 2004; Schmolck et al., 2000). This finding was nicely shown in a study examining flashbulb memory for a consequential event in recent history: the terrorist attacks on the United States on September 11, 2001 (Talarico & Rubin, 2003). Participants were asked to describe and rate the emotional event immediately after it occurred along with an ordinary event that had recently occurred. They were then asked to rate these memories again 1, 6, or 32 weeks later. Flashbulb and everyday memory equally declined in consistency over time, suggesting no difference in accuracy. The only difference was that flashbulb memory continued to be rated with high vividness, belief, and accuracy, whereas such ratings declined over time for everyday memories. Another study examining the long-term retention (up to 3 years later) of flashbulb memory for the 9/11 terrorist attacks in over 3,000 participants further showed that people are less accurate in recalling their emotional reactions to flashbulb events than in recalling nonemotional details of the event (e.g., who told them about the attack; Hirst et al., 2009). Even after a substantial delay of 10 years, high confidence in flashbulb memory still persists along with continued inconsistencies, as compared to everyday memory (Hirst et al., 2015).

Neuroimaging evidence also suggests that a heightened sense of remembering has distinct neural signatures for emotional versus neutral stimuli (Sharot et al., 2004). That is, vividly remembering an emotionally negative photo engages the amygdala, which has been linked to emotional arousal and perceptual fluency, as described; whereas remembering a neutral photo engages the posterior parahippocampus, which is associated with successful memory retrieval (Slotnick & Schacter, 2004) and recognition of perceptual details (Cabeza, Rao, Wagner, Mayer, & Schacter, 2001). Having more close personal experience with a flashbulb event, such as living near versus far from where the catastrophic event took place, is further associated with a greater enhancement of amygdala activity when recollecting the flashbulb event relative to a neutral event (Sharot, Martorella, Delgado, & Phelps, 2007). Together, these data suggest that the amygdala may facilitate the heightened sense of remembering, but it is not necessarily associated with greater accuracy for emotional events.

Other researchers extended these findings by demonstrating that emotion might be influencing vividness at retrieval by enhancing the amount of detail remembered. Specifically, participants performed better for emotional material than neutral material on tasks where they were asked to distinguish between previously seen images and similar images that were never seen before (e.g., same verbal label or category; Kensinger, Garoff-Eaton, & Schacter, 2006). When performing other types of categorization tasks, such as distinguishing between previously seen objects versus objects that had only been imagined (Kensinger & Schacter, 2005) or distinguishing between stimuli in an emotional context versus neutral context (Smith, Stephan, Rugg, & Dolan, 2006), correct attributions for emotional information were associated with greater activity in the amygdala, hippocampus, and orbitofrontal cortex, suggesting a role for these regions in aiding the remembrance of emotional events with enhanced detail.

In summary, emotion affects what reaches our awareness, ultimately shaping what is stored and later can be remembered for future survival. (For review, see Kensinger & Kark, 2017.) This is accomplished through the critical involvement of MTL regions, such as the amygdala and hippocampus, and their reciprocal interactions with each other and with other brain regions (e.g., visual processing stream). However, emotion also influences higher-order cognitive functioning (e.g., decision making) by helping us predict how particular options might make us feel in the future as well as altering choice across a variety of contexts.

INFLUENCE OF EMOTION ON COGNITION: DECISION MAKING

Affective Forecasting and Impact Bias

When making an important decision (e.g., where to live), we try to imagine how good or how bad each option will make us feel. This act—predicting our future feelings—is called affective forecasting (Wilson & Gilbert, 2003). We tend to make affective predictions about the specific emotions we might feel (e.g., joy or fear) and the duration and intensity of such emotions, which can help us decide what to do.

One key source of affective information when forecasting our future comes from recalling relevant memories from our past. Autobiographical memories are reconstructed at the time of retrieval rather than being an exact replica of the original experience (Tulving, 2002). Therefore, it comes as no surprise that remembering the past and imagining the future are thought to have overlapping neural circuitry, particularly recruitment of the medial prefrontal cortex, lateral and medial parietal cortex, and MTL including the hippocampus (Addis, Wong, & Schacter, 2007; Sharot, Riccardi, Raio, & Phelps, 2007; Szpunar & Schacter, 2017). This is adaptive when choosing between different options, because we can search for specific experiences in our past that are similar to the choice at hand. Even when imagining novel events (e.g., visiting Paris), we use different fragments of our past to piece together our best prediction of what it might be like (e.g., memories of a trip to Europe or eating French food).

Importantly, reminiscing about autobiographical memories often can transport us back in time to how we felt then (Bower, 1981; Westermann, Spies, Stahl, & Hesse, 1996), such as feeling joy when remembering the birth of a child or feeling sadness when remembering a falling out with a close friend, which is also supported by neuroimaging evidence (Damasio et al., 2000; Markowitsch, Vandekerckhove, Lanfermann, & Russ, 2003; Sheldon et al., 2017; Speer, Bhanji, & Delgado, 2014; Svoboda, McKinnon, & Levine, 2006). Retrieval often occurs involuntarily, which can be effortless in a sense (Berntsen & Hall, 2004). We can then use our momentary feelings triggered by memories to guide decision making and make judgments about the world around us (Clore, Gasper, & Garvin, 2001), which is especially beneficial in absence of other information (Suddendorf, Addis, & Corballis, 2009).

Although people do not forget *what* made them feel good or bad in the past, they do have trouble remembering exactly *how* good or bad it made them feel (e.g., Thomas & Diener, 1990). This fact often leads to an extreme prediction about our future feelings, especially for negative events, which is called the impact bias (Wilson & Gilbert, 2003). For instance, people often predict that they will feel worse than they actually end up feeling, even when the particular instance is similar to something they have already experienced in the past (Miloyan & Suddendorf, 2015).

In an elegant study, Morewedge and colleagues (2005) examined whether people tend to rely on memorable yet atypical instances as a basis for predicting their future feelings. To test this, they asked participants to describe either "any" instance of an event (e.g., a time you missed the train) or a particularly atypical instance of an event (e.g., the worst time you missed the train). Interestingly, when asked to rate past feelings, we might think that atypical experiences would have stuck out in the ratings, but, in reality, an atypical instance and a typical instance were rated with the same degree of negativity, suggesting that typical recallers may have recalled the worst instance without realizing it. When asked how the same setback (missing the train) would make them feel in the future, typical recallers were more likely to predict feeling very unhappy, because they may have generalized the worst instance as being typical or commonplace. The same finding was observed when people were asked to recall "any" instance of a positive event as well. Although atypical events may be the first to come to mind in everyday life, these findings highlight how the ease at which we recall atypical events may lead us to rely on such unrepresentative events to forecast our future feelings. Thus, impact bias can be problematic for our decision making as it might demotivate us from succeeding at a challenge because we exaggerate an overly negative reaction to the future event or by even motivating us toward a challenge that we cannot quite handle because we exaggerate an overly positive reaction to the future event.

Impact bias also can stem from the fact that we have the strongest memory for both the peak intensity and the tail end of a specific event, whereas the other details surrounding the event dissipate over time (Garbinsky, Morewedge, & Shiv, 2014). When remembering a loss, for instance, we may vividly remember the most distressing aspect of the loss (e.g., breakup with a romantic partner) while perhaps failing to remember crucial details that reduced our distress more quickly than we remember (e.g., social support from friends, finding another romantic partner). This may lead us to have a poor sense of how well we can cope with tragedy or setbacks, which is called immune neglect (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). When imagining a future event, if one discounts how well one rationalized a past failure or forgets that other positive events occurred in close proximity to a past negative event, one may be less likely to realize that a repeat of this failure will not actually feel so bad (Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000).

On the other side of the spectrum, we exaggerate our positive emotions from the past as well. This type of overestimation is associated with increased motivation on tasks that require effort. In several studies, half of the participants were explicitly asked to think about how happy it would make them feel to beat an opponent on an effortful task, whereas the other half were not asked to predict their happiness associated with winning (Greitemeyer, 2009; Morewedge & Buechel, 2013). Individuals who made the most extreme positive predictions expended the most effort on the task, such as making more button presses (Morewedge & Buechel, 2013) or persisting longer on an intelligence test (Greitemeyer, 2009), which resulted in the best performance. This is also true in a collaborative context. Individuals who overestimated how positive they would feel about a future success had an easier time recruiting collaborators to help them (von Hippel & Trivers, 2011). Additionally, it has been proposed that individuals may even purposely overestimate the affective impact of a particular event to increase motivation toward a goal. Participants who believed they had control over their performance exerted more effort than individuals who had no control (Morewedge & Buechel, 2013). This finding was true when individuals had actual control over their performance and even when they just had the perception of control.

However, it is important to note that extreme forecasts of future positive feelings also can motivate individuals to persist longer on unsolvable tasks, highlighting the fact that extreme positive predictions are not always beneficial (Greitemeyer, 2009). Taken together, research on the impact bias suggests that whether exaggerated feelings are beneficial or detrimental to motivation and decision making is dependent on the context, and that the impact bias is often detrimental when such feelings become too extreme and go unacknowledged.

Temporal Discounting

Humans have the tendency to prefer smaller, immediate rewards over larger, future rewards, which is known as temporal discounting (Ainslie, 1975). Discounting can be adaptive or maladaptive depending on the context. Acting opportunistically is adaptive to the extent that rewards are useful only if they actually occur, which is more likely now rather than later (Frederick, Loewenstein, & O'Donoghue, 2002). Impulsivity in excess can lead to maladaptive decision making, such as gambling, drug abuse, or obesity, whereas delaying reward gratification actually would be more beneficial in the long term (e.g., financial stability, sobriety; Kirby, Petry, & Bickel, 1999).

A common paradigm for studying temporal discounting is an intertemporal choice task, which asks participants to choose between different monetary rewards that are available at different points in time (e.g., $10 today vs. $20 next week). It was initially suggested that people prefer smaller, immediate rewards because they elicit greater emotional responses (Laibson, 1997), which was corroborated by functional magnetic resonance imaging evidence demonstrating that immediate rewards relative to delayed rewards show greater activity in regions previously implicated in reward processing (e.g., medial orbitofrontal cortex [OFC] and ventral striatum; McClure, Laibson, Loewenstein, & Cohen, 2004). Yet other studies yield conflicting findings, suggesting a more complex relationship. For instance, reward-related activity (e.g., striatum and ventromedial PFC [VMPFC]) correlated with the subjective value of both immediate and delayed monetary rewards (Kable & Glimcher, 2010), and lesions to the OFC actually increase temporal discounting rates (Sellitto, Ciaramelli, & di Pellegrino, 2010).

More recently, researchers have measured arousal responses during this paradigm to examine how emotion modulates intertemporal choices. In a clever set of studies utilizing pupil dilation as a measure of emotional arousal, Lempert and colleagues (2015) found that greater emotional arousal was associated with reward outcomes that were better than expected. This occurred for both immediate and delayed rewards, suggesting that emotion influences intertemporal choice in a context-dependent manner. (For review, see Lempert & Phelps, 2015.) Other researchers have manipulated emotion associated with choice to subsequently change choice behavior. Increasing vividness and emotional intensity by asking participants to imagine specific ways they would spend the delayed future reward can reduce temporal discounting and was associated with increased connectivity between the hippocampus and the VMPFC (Benoit, Gilbert, & Burgess, 2011), which are regions previously implicated in memory and future imagination (Schacter & Addis, 2007) as well as optimism (Sharot, Riccardi, et al., 2007).

Observations of reduced temporal discounting also occur when manipulating emotion not relevant to the choice at hand. For example, asking participants to engage in positive future imagination about the self before making a choice (Liu, Feng, Chen, & Li, 2013), reminiscing about positive autobiographical memories (Lempert, Speer, Delgado & Phelps, 2017), or remembering times that they felt grateful also can lead to more patient choices (DeSteno, Li, Dickens, & Lerner, 2014; Dickens & DeSteno, 2016). In contrast, inducing acute stress, a negative affective state, enhances temporal discounting and leads to more impulsive choices (Kimura et al., 2013), although this effect may be mediated by individual differences in perceived stress (Lempert, Porcelli, Delgado, & Tricomi, 2012).

Risk Taking and Loss Aversion

Risk taking—engaging in behaviors with uncertain outcomes—can be adaptive for learning new things or acquiring unforeseen opportunities. It also can be problematic, however, if risk taking develops into an addiction or leads to detrimental consequences driven by the prospect of an immediate reward (e.g., large monetary loss from gambling). A common way of measuring risky decision making in an experimental setting is by asking participants to make choices between options with differing probabilities of gains or losses associated with them.

One such paradigm is the Iowa gambling task (Bechara, Damasio, Damasio, & Anderson, 1994). In this task, participants are asked to select cards from four decks, which will result in monetary gains or losses. Unbeknownst to participants, two decks are safe (small gains and losses) and two decks are risky (larger gains but also occasional large losses). Over time, healthy participants shift their preference to the safe decks, as they learn that they are more beneficial in the long term. Choosing the risky decks was associated with an enhanced anticipatory skin conductance response (SCR)—a measure of autonomic nervous system arousal (Bechara, Damasio, Tranel, & Damasio, 1997). In stark contrast, patients with amygdala or OFC lesions do not show a preference shift to the safe decks and lack the anticipatory SCR response to risky choices (Bechara, Damasio, Damasio, & Lee, 1999). These researchers interpreted their findings as suggesting that anticipatory arousal responses may be deterring participants from choosing risky options. However, others have challenged this claim, given that avoidance and autonomic responses emerge from different neural circuitry (LeDoux & Gorman, 2001). In other decision-making contexts, such as those that mimic gambling (e.g., pay-to-play paradigms), enhanced SCR and positive emotion ratings (excitement) in response to monetary gains relative to losses predicted gambling propensity (Lole, Gonsalvez, Blaszczynski, & Clarke, 2012). In this way, emotion appears to be increasing risky decision making.

Another task that more closely captures real-world risk behavior is the Balloon Analogue Risk Task (BART; Lejuez et al., 2002). In this task, participants are asked to pump up a balloon on a screen via button presses. Each time the balloon inflates, a greater potential monetary reward is added to the balloon. But inflating the balloon also brings greater risk because all the potential earnings are lost if the balloon pops. Only when participants decide to stop inflating the balloon before it pops are they able to collect the monetary reward associated with it. Therefore, greater risk seeking is synonymous with giving unexploded balloons a greater number of pumps. Risk seeking during the BART is correlated with self-report measures of sensation seeking and impulsivity (Hunt, Hopko, Bare, Lejuez, & Robinson, 2005; Lauriola, Panno, Levin, & Lejuez, 2014; Lejuez et al., 2002) and engages the striatum, anterior insula, and dorsolateral prefrontal cortex (DLPFC) (Rao, Korczykowski, Pluta, Hoang, & Detre, 2008) as well as the VMPFC, which is consistent with reward seeking (Fukunaga, Brown, & Bogg, 2012).

Situated within the risky decision-making framework is loss aversion. Loss aversion is the preference of individuals to avoid losses over and above maximizing gains (Kahneman & Tversky, 1979). In the context of financial decision making, for instance, people will turn down bets unless they can potentially win twice as much money as they can potentially lose, on average (Kahneman, 2003). This may occur because people expect losses to have a stronger hedonic impact than gains (i.e., impact bias; Kermer, Driver-Linn, Wilson, & Gilbert, 2006).

Despite the general finding that people are loss averse, individuals widely vary in how loss averse they are. When utilizing a task that could distinguish loss aversion from risk sensitivity, for instance, greater loss aversion was associated with greater arousal (SCR; Sokol-Hessner et al., 2009) and enhanced amygdala activity in response to losses relative to gains (Peter Sokol-Hessner, Camerer, & Phelps, 2013). Importantly, there was no relationship with risk sensitivity, highlighting that loss aversion and risk sensitivity may be independently influenced by different contexts and reward magnitudes.

Loss aversion is not necessarily a bad thing, given that we face countless decisions in our everyday lives that deserve cautiousness. In a health context, individuals who needed treatment for an illness and overestimated how negatively the illness would make them feel were more motivated to seek treatment (Buick & Petrie, 2002). In this case, exaggerated fear actually increased their likelihood of a positive outcome, as long as the treatment was not more life threatening or dangerous than the illness itself (Smith et al., 2008). Additional evidence of the benefit of loss aversion comes from research on patients with amygdala damage (Gupta, Koscik, Bechara, & Tranel, 2011).

Because the amygdala is a key region for experiencing fear, these individuals do not show loss aversion and thus perform very poorly on monetary decision-making tasks (De Martino, Camerer, & Adolphs, 2010). In the real world, this could result in detrimental consequences, such as financial issues and gambling problems.

Various studies also have assessed how incidental emotions, such as mood state and acute stress, impact risky decision making, with mixed results. For instance, studies examining how different moods modulate risky choices have found that anger leads to more risk seeking, whereas fear and anxiety lead to less risk seeking (Lerner & Keltner, 2001; Raghunathan & Pham, 1999) and positive mood leads to greater loss aversion (Isen, Nygren, & Ashby, 1988). With respect to the application of acute stress, reports of risk aversion (Pabst, Brand, & Wolf, 2013) and risk seeking (Starcke, Wolf, Markowitsch, & Brand, 2008) have been observed, with potential differences also stemming from how decisions are presented; participants can at times be risk averse when choosing between potential gains but risk seeking when choosing between potential losses (Porcelli & Delgado, 2009).

There is also evidence that risky decision making under stress may be modulated by gender (Lighthall, Mather, & Gorlick, 2009; Preston, Buchanan, Stansfield, & Bechara, 2007). Using the BART described earlier, Lighthall and colleagues (2012) found men to be more risk seeking under stress, which was linked to greater activity in the dorsal striatum and insula; women showed the opposite effect as they were less risk seeking and showed diminished activity in these same regions. As suggested by the authors, enhanced dorsal striatum and insula responses seemed to be associated with reward-motivated behavior under low risk (predictable but small rewards) for stressed males, whereas women showed a more typical reward responsiveness dip consistent with prior work (Bogdan & Pizzagalli, 2006). Together, these studies highlight the role of individual differences and level of risk across different tasks for understanding how acute stress impacts risky decision making.

Social Context

Not all decisions are made in isolation. That is, many of our choices involve other people or might impact other people in some way. Social interactions can also elicit complex emotional responses, as we might have to weigh others' motives or emotional states against our own, which ultimately shapes our decision making.

One way to examine the effect of emotion on decision making in a social context is to ask participants to make choices that would potentially affect others. For instance, one study asked participants to distribute money between themselves and a charity organization (Brosch, Coppin, Scherer, Schwartz, & Sander, 2011). On each trial they were given a different monetary amount, and could choose between being altruistic (giving up money to give to charity) or being selfish (keep money at the cost of the charity). What they found is that participants who behaved more selfishly showed enhanced activity for selfish choices in regions previously linked to processing reward value representation, such as the amygdala and ventral striatum (Delgado, Nystrom, Fissell, Noll, & Fiez, 2000).

The presence of another person can also influence decisions to share or to behave selfishly. This has been explored in the ultimatum game, which is widely used in neuroeconomic experiments (Henrich et al., 2005). In this game, across many trials participants take turns as either the proposer, who decides how much money to share with a responder, or as the responder, who decides whether to accept or reject the offer from the proposer. The proposer can share equally (50:50), keep everything for him- or herself (100:0), or make any other offer in between. Given that both participants receive nothing if the responder rejects the offer, it is beneficial to accept all offers above 0 to maximize rewards. Yet responders frequently reject offers when they are presumably perceived as unfair—that is, when the proposer only offers about a quarter or less of the total sum (Henrich et al., 2005; Thaler, 1988). Intuitively, this suggests that participants sacrifice their own personal gain to punish the proposer for making an unfair offer, which is called altruistic punishment. A greater proclivity to reject unfair offers has been further associated with enhanced SCR arousal (van't Wout, Kahn, Sanfey, & Aleman, 2006) and increased activity in the anterior insula (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003)—a region associated with proprioception or the representation of bodily changes and emotions (Critchley, Elliott, Mathias, & Dolan, 2000). Interestingly, altruistic punishment seems to be perceived as rewarding despite financial costs to the self, as it engages reward-related neural circuitry (i.e., striatum; de Quervain et al., 2004), is linked to collaborative behavior in groups (Fehr & Gächter, 2002), and may be driven by a desire to achieve equality (Fowler, Johnson, Smirnov, Fehr, & Gächter, 2005).

However when performing the ultimatum game under acute stress, stress participants were less generous than non-stressed controls and engaged in more altruistic punishment immediately after stress than they did after a delay (Vinkers et al., 2013). In other decision-making contexts, such as tasks asking participants to make moral choices about everyday moral dilemmas, stress was associated with more self-centered decisions (Starcke, Polzer, Wolf, & Brand, 2011). Interestingly, undergoing acute stress can also increase prosocial behavior, in terms of greater trust and less punishment (von Dawans, Fischbacher, Kirschbaum, Fehr, & Heinrichs, 2012). Although this seems counterintuitive, this finding suggests that social approach behavior after experiencing a stressor in everyday life may serve as an adaptive strategy for stress-buffering.

INFLUENCE OF COGNITION ON EMOTION

Cognitive Emotion Regulation

In our everyday lives, we encounter challenges and setbacks that can evoke unwanted feelings. Sometimes it is possible to change our external situation to regulate these feelings (e.g., avoiding a grumpy neighbor). When this is not possible, we must change our internal responses instead. Humans employ various cognitive strategies to regulate our emotional state (Brans, Koval, Verduyn, Lim, & Kuppens, 2013; see Chapter 15 in this volume). For instance, we might try to dampen aversive feelings altogether, known as suppression, or we could limit our attentional resources to an emotionally aversive stimulus in favor of something better, known as distraction (Ochsner & Gross, 2005). However, the regulatory strategy most widely studied is cognitive change, specifically cognitive reappraisal (Buhle et al., 2014; Webb, Miles, & Sheeran, 2012).

Cognitive reappraisal involves altering the way we think about an emotional stimulus to change how we feel about it (Gross, 2002). In an experimental setting, participants are typically asked to re-interpret a stimulus (e.g., images, film clips, words) to reduce its emotional impact, such as reframing the meaning of a situation or our emotional response to it. For instance, anxious arousal before public speaking could be reframed as feelings of excitement. Countless studies have shown this strategy to be highly efficacious in reducing negative emotion via subjective self-report, or increasing positive emotion depending on the regulatory goal (e.g., Ray, McRae, Ochsner, & Gross, 2010; for review see Webb et al., 2012). There is also evidence that cognitive reappraisal can diminish physiological arousal, although the findings have been mixed (Hofmann, Heering, Sawyer, & Asnaani, 2009; Kim & Hamann, 2012). Importantly, cognitive reappraisal is central to therapeutic techniques, such as cognitive behavioral therapy (Beck, 2011); and its frequent use in everyday life is linked to fewer depressive symptoms and more positive emotionality (Gross & John, 2003), underscoring this strategy's significance for psychological well-being.

An emerging literature has begun to explore the neural mechanisms underlying cognitive emotion regulation. For instance, observations from neuroimaging studies show a reduction in amygdala activity when down-regulating negative emotion using cognitive reappraisal, along with increases in prefrontal regions linked to cognitive control and response inhibition, such as the DLPFC (e.g., Goldin, McRae, Ramel, & Gross, 2008; Kanske, Heissler, Schönfelder, Bongers, & Wessa, 2011; Ochsner et al., 2004). Although there are no direct connections between the amygdala and DLPFC, there is evidence to suggest that other prefrontal regions mediate this effect such as the ventromedial prefrontal cortex (VMPFC) or ventrolateral prefrontal cortex (VLPFC; Ochsner, Silvers, & Buhle, 2012; Wager, Davidson, Hughes, Lindquist, & Ochsner, 2008). A recent study further demonstrated that the strength of DLPFC–VLPFC/inferior frontal gyrus connectivity was associated with cognitive reappraisal success (Morawetz, Bode, Baudewig, Kirilina, & Heekeren, 2016). Interestingly, a similar prefrontal circuitry is engaged when down-regulating unwanted positive emotional responses, such as reward-related responses in the striatum elicited by anticipated monetary reward (Delgado, Gillis, & Phelps, 2008) or craving (Kober, Kross, Mischel, Hart, & Ochsner, 2010).

It is therefore intuitive that the ability to change our feelings can also change what we remember and the choices we make. Although there have been few investigations thus far, cognitive emotion regulation seems to positively impact our memory. For instance, individuals who reported using the cognitive reappraisal strategy more frequently preceding an exam actually remembered the stressful experience of exam taking to be more positive and pleasant than individuals who utilized the strategies of distraction and suppression (Ahn et al., 2015; Levine, Schmidt, Kang, & Tinti, 2012). This result provides initial evidence that changing our thoughts to change our emotions can have a lasting change on our memory for a particular event.

In the context of decision making, utilizing cognitive reappraisal can modulate choices to engage in risk taking. For example, asking participants to reappraise the significance of a choice before making it reduced both physiological arousal measured via SCR and amygdala responses to losses (Sokol-Hessner et al., 2009). In a task where participants regulated or did not regulate their emotions before choosing between a safe and a risky monetary lottery, an imagery-based reappraisal strategy (imagining a calming scene) led to fewer risky decisions (choosing the safe monetary lottery option more frequently) (Martin & Delgado, 2011). Successful emotion regulation was further associated with attenuated responses in the ventral striatum, suggesting greater cognitive control. In a social context, asking participants to cognitively reinterpret the intentions of another player in the ultimatum game led to fewer rejections of unfair offers when in the responder role and making fewer unfair offers when in the proposer role (van't Wout, Chang, & Sanfey, 2010). Thus, cognitive emotion regulation strategies not only help us reframe our experiences to change unwanted feelings; it can have a powerful impact on our subsequent decisions as well.

Mood Disorders

Given that balanced interactions between cognition and emotion contribute to healthy psychological functioning, understanding how and why these interactions go awry provides critical insight into vulnerabilities to psychiatric disorders. One psychiatric disorder characterized by disruptions to both cognition and emotion that comes to mind is depression. A key feature of depression is persistent rumination about past negative events (Patel et al., 2007; Watkins, 2008). Rumination seems to be uncontrollable and evokes unpleasant feelings. Although prior work has shown that reminiscing about positive memories is intrinsically rewarding by engaging reward-related neural circuitry, leading individuals to forgo monetary rewards for the opportunity to reexperience them (Speer et al., 2014), and can dampen the physiological stress response (i.e., cortisol; Speer & Delgado, 2017), this is not the case in depression (Chen, Takahashi, & Yang, 2015; Dillon, 2015). Depressed individuals show deficits in recalling specific memories from their past (Crane, Barnhofer, Visser, Nightingale, & Williams, 2007; Raes, Schoofs, Grif, & Hermans, 2012) and have trouble recalling positive memories in particular (Young, Bellgowan, Bodurka, & Drevets, 2013).

When describing past events, for instance, individuals with major depressive disorder and those in remission used significantly more negative attributes to describe meaningful past events than their nondepressed counterparts (Dalgleish, Hill, Golden, Morant, & Dunn, 2011). In another study, when asked to predict feelings about an upcoming Valentine's Day, individuals with greater depressive symptoms had more negatively biased predictions in the context of both a positive future event (i.e., having a date) and a negative future event (i.e., not having a date) compared to their actual feelings when Valentine's Day occurred (Hoerger, Quirk, Chapman, & Duberstein, 2012). Yet this was not found for anxiety or hypomania symptoms (i.e., mild positive mania). Altogether these findings suggest that depression may make it even less likely for people to recall the upside to a negative past situation, making it particularly challenging to employ cognitive emotion regulation strategies during depressive episodes (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007).

Neuroimaging studies provide converging evidence of cognitive emotion regulation deficits in depression. Specifically, depressed individuals show reduced DLPFC and/or VLPFC activity when attempting to down-regulate negative feelings. (For review, see Rive et al., 2013.) In the context of positive emotion, depressed individuals have difficulty sustaining both positive feelings and ventral striatal responses when attempting to up-regulate positive emotion in response to positive stimuli (Heller et al., 2009). This deficit is also accompanied by diminished prefrontal functioning, which is remedied with antidepressant treatment (Heller et al., 2013).

Because depression is associated with persistent and sometimes exaggerated negative feelings, it has the potential to create a grim outlook of the future that might lead to decreased goal-directed behavior and maladaptive decision making (Wenze, Gunthert, & German, 2012). This bias fits well with the finding that depressed individuals often experience anhedonia, which is diminished pleasure for things that were once enjoyable (Chentsova-Dutton & Hanley, 2010). If people believe that nothing good ever happened to them in the past, then there is little reason to be proactive about the future. For depressed individuals, this negative outlook inadvertently decreases the likelihood of reaching goals that presumably would generate positive emotions or increase self-efficacy, both of which would help stop this detrimental cycle.

Curiously, anxiety does not always result in the same biases in decision making as depression, despite having shared negative symptoms, such as rumination. This disparity may be because anxiety often is associated with increased worry that leads to avoidant behavior (N. Schmidt, 1994) rather than producing an overall negative outlook on one's past, as exhibited in depression. In this way, anxiety may motivate individuals to be overly cautious in terms of risk taking but perhaps does not necessarily stop anxious individuals from being proactive about their goals in a positive domain (e.g., excessive studying to guarantee success on an upcoming exam). Indeed, there is evidence suggesting that depression, but not anxiety, is associated with pessimistic predictions of future mood both in short-term and long-term thinking (Wenze, Gunthert, Ahrens, & Bos, 2013). The difference here may lie in how depression results in negatively biased autobiographical memory whereas this is not the case in anxiety. Due to fear of threat in the environment, anxiety actually may serve to lessen risk taking and increase goal-directed behavior, such as reducing temporal discounting of future rewards (Rounds, Beck, & Grant, 2007).

CONCLUSION

This chapter provided a review and synthesis of emotion–cognition interactions derived from neural and psychological experiments examining memory, decision making, and cognitive regulation. First, we emphasized how emotional arousal influences the cognitive processing of memory via the amygdala, which helps detect emotionally relevant stimuli in our environment and facilitates the transfer of important emotional information into storage for later retrieval. Heightened emotionality can further enhance our subjective sense of remembering, leading to a recollective experience with tremendous vividness and detail.

Second, we emphasized how emotion influences the cognitive processing of decision making, highlighting how we use remembered feelings from our past to predict future feelings associated with different choice options, which, similar to memory, engages the MTL. We also examined how emotion can sway decisions during intertemporal choice, under risky circumstances or even in the presence of another person. Across various decision-making contexts, the VMPFC and striatum were the most consistently activated, which is unsurprising given their role in coding subjective value and reward representation.

Finally, we emphasized how cognition influences emotion processing, with a specific focus on changing how we think to change how we feel (e.g., cognitive emotion regulation), which typically engages prefrontal regions implicated in cognitive control (e.g., DLPFC) thought to exert an influence over positive or negative emotional responses, depending on the regulatory goal.

Although emotion and cognition certainly have unique contributions, their underlying neural and psychological mechanisms are often inextricably linked. However, our knowledge of the reciprocal relationship between emotion and cognition still is limited. An intriguing question is how different aspects of emotion, such as valence versus arousal, interact to influence neural signals of cognition. Another avenue for future inquiry might be to examine the neural circuitry associated with how these interactions emerge in development and change across the life span. A better understanding of emotion–cognition interactions will provide critical insight into healthy psychological functioning while also shedding light on individual-level susceptibility to psychiatric disorders, such as depression and anxiety.

REFERENCES

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363–1377.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*(6507), 669–672. http://doi.org/10.1038/372669a0.

Ahn, H. M., Kim, S. A., Hwang, I. J., Jeong, J. W., Kim, H. T., Hamann, S., & Kim, S. H. (2015). The effect of cognitive reappraisal on long-term emotional experience and emotional memory. *Journal of Neuropsychology*, *9*, 64–76. http://doi.org/10.1111/jnp.12035.

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463–496. http://doi.org/10.1037/h0076860.

Amaral, D. G., Behniea, H., & Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience*, *118*(4), 1099–1120. http://doi.org/10.1016/S0306-4522(02)01001-1.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General*, *134*(2), 258–281. http://doi.org/10.1037/0096-3445.134.2.258.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. E. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience*, *23*(13), 5627–33. http://www.ncbi.nlm.nih.gov/pubmed/12843265.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*(6835), 305–309. http://doi.org/10.1038/35077083.

Ballarini, F., Moncada, D., Martinez, M. C., Alen, N., & Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(34), 14599–14604. http://doi.org/10.1073/pnas.0907078106.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. http://doi.org/10.1037/0033-2909.133.1.1.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7–15. http://doi.org/10.1016/0010-0277(94)90018-3.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, *19*(13), 5473–5481. http://www.ncbi.nlm.nih.gov/pubmed/10377356.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. (1997). Deciding advantageous before knowing the advantageous strategy. *Science*, *275*, 1293–1294. http://doi.org/10.1126/science.275.5304.1293.

Beck, J. S. (2011). *Cognitive Behavior Therapy: Basics and Beyond*, 2nd ed. New York, NY: Guilford Press.

Benoit, R. G., Gilbert, S. J., & Burgess, P. W. (2011). A neural mechanism mediating the impact of episodic prospection on farsighted decisions. *Journal of Neuroscience*, *31*(18), 6771–6779. http://doi.org/10.1523/JNEUROSCI.6559-10.2011.

Berntsen, D., & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition*, *32*(5), 789–803.

Bogdan, R., & Pizzagalli, D. A. (2006). Acute stress reduces reward responsiveness: Implications for depression. *Biological Psychiatry*, *60*(10), 1147–1154. http://doi.org/10.1016/j.biopsych.2006.03.037.

Bonnet, L., Comte, A., Tatu, L., Millot, J.-L., Moulin, T., & Medeiros de Bustos, E. (2015). The role of the amygdala in the perception of positive emotions: an "intensity detector." *Frontiers in Behavioral Neuroscience*, *9*, 178. http://doi.org/10.3389/fnbeh.2015.00178.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*(2), 129–148.

Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). The regulation of negative and positive affect in daily life. *Emotion*, *13*(5), 926–939. http://doi.org/10.1037/a0032400.

Brosch, T., Coppin, G., Scherer, K. R., Schwartz, S., & Sander, D. (2011). Generating value(s): Psychological value hierarchies reflect context-dependent sensitivity of the reward system. *Social Neuroscience*, *6*(2), 198–208. http://doi.org/10.1080/17470919.2010.506754.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*(1), 73–99. http://doi.org/10.1016/0010-0277(77)90018-X.

Buhle, J. T., Silvers, J., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., . . . Ochsner, K. N. (2014). Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cerebral Cortex*, *24*(11), 2981–2990. http://doi.org/10.1093/cercor/bht154.

Buick, D. L., & Petrie, K. J. (2002). "I know just how you feel ": the validity of healthy women's perceptions of breast-cancer patients receiving treatment. *Journal of Applied Social Psychology*, *32*(1), 110–123.

Cabanac, M. (2002). What is emotion? *Behavioural Processes*, *60*, 69–83.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences*, *98*(8), 4805–4810. http://doi.org/10.1073/pnas.081082698.

Cahill, L., & Alkire, M. T. (2003). Epinephrine enhancement of human memory consolidation: interaction with arousal at encoding. *Neurobiology of Learning and Memory*, *79*(2), 194–8. http://www.ncbi.nlm.nih.gov/pubmed/12591227.

Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: interaction with the degree of arousal at encoding. *Learning & Memory*, *10*(4), 270–274. http://doi.org/10.1101/lm.62403.

Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Keator, D., . . . McGaugh, J. L. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(15), 8016–21. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=38867&tool=pmcentrez&rendertype=abstract.

Cahill, L., Prins, B., Weber, M., & McGaugh, J. L. (1994). β-adrenergic activation and memory for emotional events. *Nature*, *371*(6499), 702–704. http://doi.org/10.1038/371702a0.

Chen, C., Takahashi, T., & Yang, S. (2015). Remembrance of happy things past: positive autobiographical memories are intrinsically rewarding and valuable, but not in depression. *Frontiers in Psychology*, *6*, 222. http://doi.org/10.3389/fpsyg.2015.00222.

Chentsova-Dutton, Y., & Hanley, K. (2010). The effects of anhedonia and depression on hedonic responses. *Psychiatry Research*, *179*(2), 176–180. http://doi.org/10.1016/j.psychres.2009.06.013.

Clore, G. L., Gasper, K., & Garvin, E. (2001). Affect as information. In J. P. Forgas, (Ed.). Handbook of affect and social cognition (pp. 121–144). Mahwah, NJ.: Erlbaum.

Cohen, J. D. (2005). The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives*, *19*(4), 3–24.

Crane, C., Barnhofer, T., Visser, C., Nightingale, H., & Williams, J. M. G. (2007). The effects of analytical and experiential rumination on autobiographical memory specificity in individuals with a history of major depression. *Behaviour Research and Therapy*, *45*(12), 3077–3087. http://doi.org/10.1016/j.brat.2007.05.009.

Critchley, H. D., Elliott, R., Mathias, C. J., & Dolan, R. J. (2000). Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *Journal of Neuroscience*, *20*(8), 3033–3040. http://www.ncbi.nlm.nih.gov/pubmed/10751455.

Cunningham, W. A., & Brosch, T. (2012). Motivational salience: amygdala tuning from traits, needs, values, and goals. *Current Directions in Psychological Science*, *21*(1), 54–59. http://doi.org/10.1177/0963721411430832.

Dalgleish, T., Hill, E., Golden, A. J., Morant, N., & Dunn, B. D. (2011). The structure of past and future lives in depression. *Journal of Abnormal Psychology*, *120*(1), 1–15. http://doi.org/10.1037/a0020797.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York, NY: Grosset/Putnam.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-enerated emotions. *Nature Neuroscience*, *3*(10), 1049–1056. http://doi.org/10.1038/79871.

Delgado, M. R., Gillis, M. M., & Phelps, E. A. (2008). Regulating the expectation of reward via cognitive strategies. *Nature Neuroscience*, *11*(8), 880–881. http://doi.org/10.1038/nn.2141.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, *84*(6), 3072–3077. http://www.ncbi.nlm.nih.gov/pubmed/11110834.

De Martino, B., Camerer, C. F., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 3788–3792. http://doi.org/10.1073/pnas.0910230107.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254. http://doi.org/10.1126/science.1100735.

DeSteno, D., Li, Y., Dickens, L., & Lerner, J. S. (2014). Gratitude: a tool for reducing economic impatience. *Psychological Science*, *25*(6), 1262–1267. http://doi.org/10.1177/0956797614529979.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology*, *12*(4), 541–551. http://doi.org/10.1080/095414400750050222.

Dickens, L., & DeSteno, D. (2016). The grateful are patient: heightened daily gratitude is associated with attenuated temporal discounting. *Emotion*, *16*(4), 421–425. http://doi.org/10.1037/emo0000176.

Dillon, D. G. (2015). The neuroscience of positive memory deficits in depression. *Frontiers in Psychology*, *6*, 1295. http://doi.org/10.3389/fpsyg.2015.01295.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*(5), 855–863. http://www.ncbi.nlm.nih.gov/pubmed/15182723

Duncan, C. P. (1949). The retroactive effect of electroshock on learning. *Journal of Comparative and Physiological Psychology*, *42*(1), 32–44. http://doi.org/10.1037/h0058173.

Dunsmoor, J. E., Murty, V. P., Davachi, L., & Phelps, E. A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. *Nature*, *520*, 345–348. http://doi.org/10.1038/nature14106.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. http://doi.org/10.1037/h0047707.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140. http://doi.org/10.1038/415137a.

Flexner, L. B., Flexner, J. B., & Stellar, E. (1965). Memory and cerebral protein synthesis in mice as affected by graded amounts of puromycin. *Experimental Neurology*, *13*(3), 264–272. http://doi.org/10.1016/0014-4886(65)90114-7.

Fowler, J. H., Johnson, T., Smirnov, O., Fehr, E., & Gächter, S. (2005). Human behaviour: egalitarian motive and altruistic punishment. *Nature*, *433*, E1. http://doi.org/10.1038/nature03256

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

Fukunaga, R., Brown, J. W., & Bogg, T. (2012). Decision making in the Balloon Analogue Risk Task (BART): anterior cingulate cortex signals loss aversion but not the infrequency of risky choices. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 479–490. http://doi.org/10.3758/s13415-012-0102-1.

Garbinsky, E. N., Morewedge, C. K., & Shiv, B. (2014). Interference of the end: why recency bias in memory determines when a food is consumed again. *Psychological Science*, *25*(7), 1466–14741–9. http://doi.org/10.1177/0956797614534268.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: a source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*(3), 617–638. http://doi.org/10.1037/0022-3514.75.3.617.

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*(6), 577–586. http://doi.org/10.1016/j.biopsych.2007.05.031.

Greitemeyer, T. (2009). The effect of anticipated affect on persistence and performance. *Personality and Social Psychology Bulletin*, *35*, 172–186. http://doi.org/10.1177/0146167208326124.

Gross, J. J. (2002). Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology*, *39*, 281–291. http://doi.org/10.1017.S0048577201393198.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362.

Gupta, R., Koscik, T. R., Bechara, A., & Tranel, D. (2011). The amygdala and decision-making. *Neuropsychologia*, *49*(4), 760–766. http://doi.org/10.1016/j.neuropsychologia.2010.09.029.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400. http://doi.org/10.1016/S1364-6613(00)01707-1.

Hamann, S., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, *2*(3), 289–293. http://doi.org/10.1038/6404.

Heller, A. S., Johnstone, T., Light, S. N., Peterson, M. J., Kolden, G. G., Kalin, N. H., & Davidson, R. J. (2013). Relationships between changes in sustained fronto-striatal connectivity and positive affect in major depression resulting from antidepressant treatment. *American Journal of Psychiatry*, *170*(2), 197–206. http://doi.org/10.1176/appi.ajp.2012.12010014.

Heller, A. S., Johnstone, T., Shackman, A. J., Light, S. N., Peterson, M. J., Kolden, G. G., . . . Davidson, R. J. (2009). Reduced capacity to sustain positive emotion in major depression reflects diminished maintenance of fronto-striatal brain activation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22445–22450.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., . . . Tracer, D. (2005). Models of decision-making and the coevolution of social preferences. *Behavioral and Brain Sciences*, *28*(6), 838–855. http://doi.org/10.1017/S0140525X05460148.

Hirst, W., Phelps, E. A., Buckner, R. L., Budson, A. E., Cuc, A., Gabrieli, J. D. E., . . . Schacter, D. L. (2009). Long-term memory for the terrorist attack of September 11: flashbulb memories, event memories, and the factors that influence their retention. *Journal of Experimental Psychology: General*, *138*(2), 161–176. http://doi.org/10.1037/a0015527.

Hirst, W., Phelps, E. A., Meksin, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., . . . Olsson, A.

(2015). A ten-year follow-up of a study of memory for the attack of September 11, 2001: flashbulb memories and memories for flashbulb events. *Journal of Experimental Psychology: General*, *144*(3), 1–71. http://doi.org/10.1037/xge0000055.

Hoerger, M., Quirk, S. W., Chapman, B. P., & Duberstein, P. R. (2012). Affective forecasting and self-rated symptoms of depression, anxiety, and hypomania: evidence for a dysphoric forecasting bias. *Cognition & Emotion*, *26*(6), 1098–1106. http://doi.org/10.1080/02699931.2011.631985.

Hofmann, S. G., Heering, S., Sawyer, A. T., & Asnaani, A. (2009). How to handle anxiety: the effects of reappraisal, acceptance, and suppression strategies on anxious arousal. *Behaviour Research and Therapy*, *47*(5), 389–394. http://doi.org/10.1016/j.brat.2009.02.010.

Hunt, M. K., Hopko, D. R., Bare, R., Lejuez, C. W., & Robinson, E. V. (2005). Construct validity of the Balloon Analog Risk Task (BART): associations with psychopathy and impulsivity. *Assessment*, *12*(4), 416–428. http://doi.org/10.1177/1073191105278740.

Isen, A. M., Nygren, T. E., & Ashby, F. G. (1988). Influence of positive affect on the subjective utility of gains and losses: it is just not worth the risk. *Journal of Personality and Social Psychology*, *55*(5), 710–717. http://doi.org/10.1037/0022-3514.55.5.710.

Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *Journal of Neuroscience*, *27*(33), 8877–8884. http://doi.org/10.1523/JNEUROSCI.2063-07.2007.

Kable, J. W., & Glimcher, P. W. (2010). An "as soon as possible" effect in human intertemporal decision making: behavioral evidence and neural mechanisms. *Journal of Neurophysiology*, *103*(5), 2513–31. http://doi.org/10.1152/jn.00177.2009.

Kahneman, D. (2003). Experiences of collaborative research. *American Psychologist*, *58*(9), 723–730. http://doi.org/10.1037/0003-066X.58.9.723.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47*(2), 263–292. http://doi.org/10.2307/1914185.

Kanske, P., Heissler, J., Schönfelder, S., Bongers, A., & Wessa, M. (2011). How to regulate emotion? Neural networks for reappraisal and distraction. *Cerebral Cortex*, *21*(6), 1379–1388. http://doi.org/10.1093/cercor/bhq216.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. http://doi.org/10.3758/BF03195800

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2006). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language*, *54*(1), 99–112. http://doi.org/10.1016/j.jml.2005.05.005.

Kensinger, E. A., & Kark, S. (2017). Emotion and memory. In J. T. Wixted, E. Phelps, & L. Davachi (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 1: *Learning and Memory*). Hoboken, NJ: Wiley.

Kensinger, E. A., & Schacter, D. L. (2005). Retrieving accurate and distorted memories: neuroimaging evidence for effects of emotion. *NeuroImage*, *27*(1), 167–177. http://doi.org/10.1016/j.neuroimage.2005.03.038.

Kermer, D. A., Driver-Linn, E., Wilson, D., & Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychological Science*, *17*(8), 649–654.

Kim, S. H., & Hamann, S. (2012). The effect of cognitive reappraisal on physiological reactivity and emotional memory. *International Journal of Psychophysiology*, *83*(3), 348–356. http://doi.org/10.1016/j.ijpsycho.2011.12.001.

Kimura, K., Izawa, S., Sugaya, N., Ogawa, N., Yamada, K. C., Shirotsuki, K., . . . Hasegawa, T. (2013). The biological effects of acute psychosocial stress on delay discounting. *Psychoneuroendocrinology*, *38*(10), 2300–2308. http://doi.org/10.1016/j.psyneuen.2013.04.019.

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, *128*(1), 78–87.

Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology*, *65*(2), 190–193. http://doi.org/10.1037/h0040288.

Kober, H., Kross, E. F., Mischel, W., Hart, C. L., & Ochsner, K. N. (2010). Regulation of craving by cognitive strategies in cigarette smokers. *Drug and Alcohol Dependence*, *106*(1), 52–55. http://doi.org/10.1016/j.drugalcdep.2009.07.017.

Labar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64. http://doi.org/10.1038/nrn1825.

LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: role of the medial temporal lobe in humans. *Psychological Science*, *9*(6), 490–493. http://doi.org/10.1111/1467-9280.00090.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, *112*(2), 443–477.

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2014). Individual differences in risky decision making: a meta-analysis of sensation seeking and impulsivity with the balloon analogue risk task. *Journal of Behavioral Decision Making*, *27*(1), 20–36. http://doi.org/10.1002/bdm.1784.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist*, *39*(2), 124–129. http://doi.org/10.1037/0003-066X.39.2.124.

LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(19), 2871–2878. http://doi.org/10.1073/pnas.1400335111.

LeDoux, J. E., & Gorman, J. M. (2001). A call to action: overcoming anxiety through active coping. *American Journal of Psychiatry*, *158*(12), 1953–1955.

Lee, T. H., Sakaki, M., Cheng, R., Velasco, R., & Mather, M. (2013). Emotional arousal amplifies the effects of biased competition in the brain. *Social Cognitive and Affective Neuroscience*, *9*(12), 2067–2077. http://doi.org/10.1093/scan/nsu015.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., . . . Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84. http://doi.org/10.1037/1076-898X.8.2.75.

Lempert, K. M., Glimcher, P. W., & Phelps, E. A. (2015). Emotional arousal and discount rate in intertemporal choice are reference-dependent. *Journal of Experimental Psychology: General*, *144*(2), 366–373. http://doi.org/10.1016/bs.mcb.2015.01.016.

Lempert, K. M., & Phelps, E. A. (2015). The malleability of intertemporal choice. *Trends in Cognitive Sciences*, *20*(1), 64–74. http://doi.org/10.1016/j.tics.2015.09.005.

Lempert, K. M., Speer, M. E., Delgado, M. R., & Phelps, E. A. (2017). Positive autobiographical memory retrieval reduces temporal discounting. *Social Cognitive and Affective Neuroscience*, nsx086. https://doi.org/10.1093/scan/nsx086.

Lempert, K. M., Porcelli, A. J., Delgado, M. R., & Tricomi, E. (2012). Individual differences in delay discounting under acute stress: the role of trait perceived stress. *Frontiers in Psychology*, *3*(251), 1–10.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159. http://doi.org/10.1037/0022-3514.81.1.146.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, *66*, 799–823. https://doi.org/10.1146/annurev-psych-010213-115043.

Levine, L. J., Schmidt, S., Kang, H. S., & Tinti, C. (2012). Remembering the silver lining: reappraisal and positive bias in memory for emotion. *Cognition & Emotion*, *26*(5), 871–885.

Lewis, D. J. (1979). Psychobiology of active and inactive memory. *Psychological Bulletin*, *86*(5), 1054–1083. http://doi.org/10.1037/0033-2909.86.5.1054.

Lighthall, N. R., Mather, M., & Gorlick, M. A. (2009). Acute stress increases sex differences in risk seeking in the Balloon Analogue Risk Task. *PLoS One*, *4*(7), e6002. http://doi.org/10.1371/journal.pone.0006002.

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E. Y., . . . Mather, M. (2012). Gender differences in reward-related decision processing under stress. *Social Cognitive and Affective Neuroscience*, *7*(4), 476–484. http://doi.org/10.1093/scan/nsr026.

Liu, L., Feng, T., Chen, J., & Li, H. (2013). The value of emotion: how does episodic prospection modulate delay discounting? *PLoS One*, *8*(11), e81717. http://doi.org/10.1371/journal.pone.0081717.

Lole, L., Gonsalvez, C. J., Blaszczynski, A., & Clarke, A. R. (2012). Electrodermal activity reliably captures physiological differences between wins and losses during gambling on electronic machines. *Psychophysiology*, *49*(2), 154–163. http://doi.org/10.1111/j.1469-8986.2011.01290.x.

Markowitsch, H. J., Vandekerckhove, M. M. P., Lanfermann, H., & Russ, M. O. (2003). Engagement of lateral and medial prefrontal areas in the ecphory of sad and happy autobiographical memories. *Cortex*, *39*, 643–665.

Martin, L. N., & Delgado, M. R. (2011). The influence of emotion regulation on decision-making under risk. *Journal of Cognitive Neuroscience*, *23*(9), 2569–2581. http://doi.org/10.1162/jocn.2011.21618.

Matheson, H. E., & Barsalou, L. (2017). Embodied cognition. In J. T. Wixted & S. Thompson-Schill (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 3: *Language and Thought*). Hoboken, NJ: Wiley.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503–507. http://doi.org/10.1126/science.1100907.

McGaugh, J. L. (2000). Memory—a century of consolidation. *Science*, *287*(5451), 248–251. http://doi.org/10.1126/science.287.5451.248.

McGaugh, J. L., & Roozendaal, B. (2002). Role of adrenal stress hormones in forming lasting memories in the brain. *Current Opinion in Neurobiology*, *12*(2), 205–210. http://doi.org/10.1016/S0959-4388(02)00306-9.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences*, *7*(3), 141–144. http://doi.org/10.1016/S1364-6613(03)00029-9.

Miloyan, B., & Suddendorf, T. (2015). Feelings of the future. *Trends in Cognitive Sciences*, *19*(4), 196–200. http://doi.org/10.1016/j.tics.2015.01.008.

Morawetz, C., Bode, S., Baudewig, J., Kirilina, E., & Heekeren, H. R. (2016). Changes in effective connectivity between dorsal and ventral prefrontal regions moderate emotion regulation. *Cerebral Cortex*, *26*(5), 1923–1937. http://doi.org/10.1093/cercor/bhv005.

Morewedge, C. K., & Buechel, E. C. (2013). Motivated underpinnings of the impact bias in affective forecasts. *Emotion*, *13*(6), 1023–1029. http://doi.org/10.1037/a0033797.

Morewedge, C. K., Gilbert, D. T., & Wilson, T. D. (2005). The least likely of times: how remembering the past biases forecasts of the future. *Psychological Science*, *16*(8), 626–630.

Nader, K., & Einarsson, E. O. (2010). Memory reconsolidation: an update. *Annals of the New York Academy of Sciences*, *1191*, 27–41. http://doi.org/10.1111/j.1749-6632.2010.05443.x.

Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, *406*(6797), 722–726. http://doi.org/10.1038/35021052.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: false recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and Accuracy in Recall* (pp. 9–31). Cambridge, UK: Cambridge University Press. http://doi.org/10.1017/CBO9780511664069.003.

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology:*

General, *129*(2), 242–261. http://doi.org/10.1037/0096-3445.129.2.242.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage*, *23*(2), 483–499. http://doi.org/10.1016/j.neuroimage.2004.06.030.

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*, E1–E24. http://doi.org/10.1111/j.1749-6632.2012.06751.x.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466–478. http://doi.org/10.1037/0096-3445.130.3.466.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. http://doi.org/10.1037/0033-295X.108.3.483.

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision making: a few minutes make all the difference. *Behavioural Brain Research*, *250*, 39–45. http://doi.org/10.1016/j.bbr.2013.04.046.

Patel, T., Brewin, C. R., Wheatley, J., Wells, A., Fisher, P., & Myers, S. (2007). Intrusive images and memories in major depression. *Behaviour Research and Therapy*, *45*, 2573–2580. http://doi.org/10.1016/j.brat.2007.06.004.

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and decision making: multiple modulatory neural circuits. *Annual Review of Neuroscience*, 263–290. http://doi.org/10.1146/annurev-neuro-071013-014119.

Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, *20*(3), 278–283. http://doi.org/10.1111/j.1467-9280.2009.02288.x.

Preston, S. D., Buchanan, T. W., Stansfield, R. B., & Bechara, A. (2007). Effects of anticipatory stress on decision making in a gambling task. *Behavioral Neuroscience*, *121*(2), 257–263. http://doi.org/10.1037/0735-7044.121.2.257.

Raes, F., Schoofs, H., Grif, J. W., & Hermans, D. (2012). Rumination relates to reduced autobiographical memory specificity in formerly depressed patients following a self-discrepancy challenge: the case of autobiographical memory specificity reactivity. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*, 1002–1007. http://doi.org/10.1016/j.jbtep.2012.03.003.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, *79*(1), 56–77. http://doi.org/10.1006/obhd.1999.2838.

Rao, H., Korczykowski, M., Pluta, J., Hoang, A., & Detre, J. A. (2008). Neural correlates of voluntary and involuntary risk taking in the human brain: an fMRI Study of the Balloon Analog Risk Task (BART). *NeuroImage*, *42*(2), 902–910. http://doi.org/10.1016/j.neuroimage.2008.05.046.

Ray, R. D., McRae, K., Ochsner, K. N., & Gross, J. J. (2010). Cognitive reappraisal of negative affect: converging evidence from EMG and self-report. *Emotion*, *10*(4), 587–592. http://doi.org/10.1037/a0019015.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860.

Richardson, M. P., Strange, B. A., & Dolan, R. J. (2004). Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nature Neuroscience*, *7*(3), 278–285. http://doi.org/10.1038/nn1190.

Rive, M. M., Rooijen, G. Van, Veltman, D. J., Phillips, M. L., Schene, A. H., & Ruhé,

H. G. (2013). Neural correlates of dysfunctional emotion regulation in major depressive disorder. A systematic review of neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(10), 2529–2553. http://doi.org/10.1016/j.neubiorev.2013.07.018.

Rounds, J. S., Beck, J. G., & Grant, D. M. (2007). Is the delay discounting paradigm useful in understanding social anxiety? *Behaviour Research and Therapy*, *45*, 729–735. http://doi.org/10.1016/j.brat.2006.06.007.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision making in the Ultimatum Game. *Science*, *300*(2003), 1755–1758. http://doi.org/10.1126/science.1082976.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London*. *Series B, Biological Sciences*, *362*(1481), 773–786. http://doi.org/10.1098/rstb.2007.2087.

Schiller, D., Monfils, M.-H., Raio, C. M., Johnson, D. C., Ledoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, *463*(7), 49–54.

Schmidt, N. (1994). The overprediction of fear and panic in panic disorder. *Behaviour Research and Therapy*, *32*(7), 701–707. http://doi.org/10.1016/0005-7967(94)90027-2.

Schmidt, S. R. (2004). Autobiographical memories for the September 11th attacks: reconstructive errors and emotional impairment of memory. *Memory & Cognition*, *32*(3), 443–454. http://doi.org/10.3758/BF03195837.

Schmolck, H., Buffalo, E. A., & Squire, L. R. (2000). Memory Distortions develop over time: recollections of the O.J. Simpson trial verdict after 15 and 32 Months. *Psychological Science*, *11*(1), 39–45. http://doi.org/10.1111/1467-9280.00212.

Schwabe, L., Nader, K., & Pruessner, J. C. (2014). Reconsolidation of human memory: brain mechanisms and clinical relevance. *Biological Psychiatry*, *76*(4), 274–280. http://doi.org/10.1016/j.biopsych.2014.03.008.

Sellitto, M., Ciaramelli, E., & di Pellegrino, G. (2010). Myopic discounting of future rewards after medial orbitofrontal damage in humans. *Journal of Neuroscience*, *30*(49), 16429–16436. http://doi.org/10.1523/JNEUROSCI.2516-10.2010.

Sharot, T., Delgado, M. R., & Phelps, E. A. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience*, *7*(12), 1376–80. http://doi.org/10.1038/nn1353.

Sharot, T., Martorella, E. A., Delgado, M. R., & Phelps, E. A. (2007). How personal experience modulates the neural circuitry of memories of September 11. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 389–394. http://doi.org/10.1073/pnas.0609230103.

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: disentangling the effects of attention and retention. *Cognitive, Affective & Behavioral Neuroscience*, *4*(3), 294–306. http://doi.org/10.3758/CABN.4.3.294.

Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, *450*(7166), 102–105. http://doi.org/10.1038/nature06280.

Sheldon, S., Diamond, N., Amrson, M., Palombo, D., Selarka, D., Romero, K., . . . Levine, B. (2017). Autobiographical memory: assessment and neuroimaging in health and disease. In J. T. Wixted, E. Phelps & L. Davachi (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 1: *Learning and Memory*). Hoboken, NJ: Wiley.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*(6), 664–672. http://doi.org/10.1038/nn1252.

Smith, A. P. R., Stephan, K. E., Rugg, M. D., & Dolan, R. J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron*, *49*(4), 631–638. http://doi.org/10.1016/j.neuron.2005.12.025.

Smith, D., Loewenstein, G., Jepson, C., Jankovich, A., Feldman, H., & Ubel, P. (2008). Mispredicting and misremembering: patients with renal failure overestimate improvements in quality of life after a kidney transplant. *Health Psychology*, *27*(5), 653–658. http://doi.org/10.1037/a0012647.

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience*, *8*(3), 341–350. http://doi.org/10.1093/scan/nss002.

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040. http://doi.org/10.1073/pnas.0806761106.

Speer, M. E., Bhanji, J. P., & Delgado, M. R. (2014). Savoring the past: positive memories evoke value representations in the striatum. *Neuron*, *84*(4), 847–856. http://doi.org/10.1016/j.neuron.2014.09.028.

Speer, M. E., & Delgado, M. R. (2017). Reminiscing about positive memories buffers acute stress responses. *Nature Human Behaviour*, *1*, 93.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177. http://doi.org/10.1016/j.nlm.2004.06.005

Starcke, K., Polzer, C., Wolf, O. T., & Brand, M. (2011). Does stress alter everyday moral decision-making? *Psychoneuroendocrinology*, *36*(2), 210–219. http://doi.org/10.1016/j.psyneuen.2010.07.010.

Starcke, K., Wolf, O. T., Markowitsch, H. J., & Brand, M. (2008). Anticipatory stress influences decision making under explicit risk conditions. *Behavioral Neuroscience*, *122*(6), 1352–1360. http://doi.org/10.1037/a0013281.

Strange, B. A., Henson, R. N. A., Friston, K. J., & Dolan, R. J. (2000). Brain mechanisms for detecting perceptual, semantic, and emotional deviance. *NeuroImage*, *12*, 425–433. http://doi.org/10.1006/nimg.2000.0637.

Suddendorf, T., Addis, D. R., & Corballis, M. C. (2009). Mental time travel and the shaping of the human mind. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*, 1317–1324. http://doi.org/10.1098/rstb.2008.0301.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia*, *44*(12), 2189–2208. http://doi.org/10.1016/j.neuropsychologia.2006.05.023.

Szpunar, K., & Schacter, D. (2017). Memory and future imagining. In J. T. Wixted, E. Phelps, & L. Davachi (Eds.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th ed., Vol. 1: *Learning and Memory*). Hoboken, NJ: Wiley.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, *14*(5), 455–461.

Thaler, R. (1988). Anomalies: the ultimatum game. *Journal of Economic Perspectives*, *2*(4), 195–206.

Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, *59*(2), 291–297. http://doi.org/10.1037/0022-3514.59.2.291.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology*, *53*, 1–25. http://doi.org/10.1146/annurev.psych.53.100901.135114.

van't Wout, M., Chang, L. J., & Sanfey, A. G. (2010). The influence of emotion regulation on social interactive decision-making. *Emotion*, *10*(6), 815–821. http://doi.org/10.1037/a0020069.

van't Wout, M., Kahn, R. S., Sanfey, A. G., & Aleman, A. (2006). Affective state and decision-making in the Ultimatum Game. *Experimental Brain Research*, *169*(4), 564–568. http://doi.org/10.1007/s00221-006-0346-5.

Vinkers, C. H., Zorn, J. V., Cornelisse, S., Koot, S., Houtepen, L. C., Olivier, B., . . . Joëls, M. (2013). Time-dependent changes in altruistic punishment following stress. *Psychoneuroendocrinology*, *38*(9), 1467–1475. http://doi.org/10.1016/j.psyneuen.2012.12.012.

von Dawans, B., Fischbacher, U., Kirschbaum, C., Fehr, E., & Heinrichs, M. (2012). The social dimension of stress reactivity: acute stress increases prosocial behavior in humans. *Psychological Science*, *23*(6), 651–660. http://doi.org/10.1177/0956797611431576.

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, *34*, 1–56. http://doi.org/10.1017/S0140525X10001354.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, *7*(11), 1271–1278. http://doi.org/10.1038/nn1341.

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, *59*(6), 1037–1050. http://doi.org/10.1016/j.neuron.2008.09.006.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*, 163–206. http://doi.org/10.1037/0033-2909.134.2.163.

Webb, T. L., Miles, E., & Sheeran, P. (2012). Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychological Bulletin*, *138*(4), 775–808. http://doi.org/10.1037/a0027600.

Wenze, S. J., Gunthert, K. C., Ahrens, A. H., & Bos, T. C. T. (2013). Biases in short-term mood prediction in individuals with depression and anxiety symptoms. *Individual Differences Research*, *11*(3), 91–101. http://doi.org/10.1016/j.biotechadv.2011.08.021.

Wenze, S. J., Gunthert, K. C., & German, R. E. (2012). Biases in affective forecasting and recall in individuals with depression and anxiety symptoms. *Personality and Social Psychology Bulletin*, *38*, 895–906. http://doi.org/10.1177/0146167212447242.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. *European Journal of Social Psychology*, *26*, 557–580.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, *18*(1), 411–418. http://www.ncbi.nlm.nih.gov/pubmed/9412517.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology*, *35*, 345–411. http://doi.org/10.1016/S0065-2601(03)01006-2.

Wilson, T., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: a source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *78*(5), 821–836. http://doi.org/10.1037/0022-3514.78.5.821.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. http://doi.org/10.1006/jmla.2002.2864.

Young, K. D., Bellgowan, P. S. F., Bodurka, J., & Drevets, W. C. (2013). Behavioral and neurophysiological correlates of autobiographical memory deficits in patients with depression and individuals at high risk for depression. *JAMA Psychiatry*, *70*(7), 698–708. http://doi.org/10.1001/jamapsychiatry.2013.1189.

Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, *39*(2), 117–123. http://doi.org/10.1037/0003-066X.39.2.117.

Author Index

Subject Index