

Claire M. Zedelius
Barbara C.N. Müller
Jonathan W. Schooler *Editors*

The Science of Lay Theories

How Beliefs Shape Our Cognition,
Behavior, and Health

 Springer

The Science of Lay Theories

Claire M. Zedelius · Barbara C.N. Müller
Jonathan W. Schooler
Editors

The Science of Lay Theories

How Beliefs Shape Our Cognition, Behavior,
and Health

 Springer

Editors

Claire M. Zedelius
Psychological and Brain Sciences
University of California, Santa Barbara
Santa Barbara, CA
USA

Jonathan W. Schooler
Psychological and Brain Sciences
University of California, Santa Barbara
Santa Barbara, CA
USA

Barbara C.N. Müller
Behavioural Science Institute
Radboud University
Nijmegen
The Netherlands

ISBN 978-3-319-57305-2 ISBN 978-3-319-57306-9 (eBook)
DOI 10.1007/978-3-319-57306-9

Library of Congress Control Number: 2017937928

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

Carol S. Dweck

What a thrill to see lay theories shedding light on everything from prejudice to creativity, thinking, self-regulation, health, freewill, and religion. It is very rare to say of an edited volume of scholarly chapters “I couldn’t put it down!” Yet that was the case with this book. It is not just that I have worked in this field for many years, but rather, with every chapter I felt I was gaining new insights into what, deep down, people really believe and how these beliefs influence their lives.

Lay theories took a while to capture the imagination of a wide swath of researchers. The cognitive revolution that began in the late 1960s shone a spotlight on thought processes, but not on beliefs or lay theories. Even in social psychology, where construals and interpretations became popular (as in attribution theory), little attention was paid to the underlying beliefs or lay theories that fostered these construals or interpretations in the first place. Researchers did not ask the deeper “Why?” And the deeper why’s are these fundamental assumptions people make about themselves and their worlds. You can call them lay theories, mindsets, world assumptions, mental models, but they are all about people’s fundamental understanding about the nature and workings of the people, things, and phenomena in their worlds.

In my own work, I came to lay theories by continually asking why. At first, I found that children’s attributions predicted their responses to failure. But I wondered *why* children with relatively equal ability would have such different interpretations of failure (with some blaming their ability and others focusing on their effort or strategies). So my colleagues and I started studying achievement goals and we found some answers there. But I still wondered *why*. Why would children of pretty equal ability have such different goals? That is when we discovered that lay theories of intelligence were at the heart of it all. Those who believed their intelligence was fixed, as opposed to developable, chose different goals and made different attributions in the face of difficulty. This is how our research on lay theories of intelligence was born.

Other researchers, at the same time, were also exploring the power of lay theories. Here are just a few examples. Melvin Lerner examined the impact of just world beliefs and what people will do to maintain their faith in that world. In a

related vein, Ronnie Janoff-Bulman examined people's assumptions about the safety and fairness of their world and studied the consequences of having those assumptions shattered. Attachment theorists were identifying people's working models of attachment—how relationships work and what you can expect from them. Even Piaget, the consummate theorist of pure logic and cognition, began to believe that people's world views might be as important as their logical thinking.

This book breaks open the field of lay theories and puts it in a much larger perspective. The chapters show how the field of lay theories has burgeoned and come to fruition. Interestingly, many of the research strands have emerged independently and are being brought together for the first time in this book.

Each chapter asks important questions about lay theories and offers intriguing and sometimes surprising answers. The chapters in the first section ask about the origins and nature of lay theories.

- Where do all these lay theories come from? Are they deeply embedded in human psychology or do they arise from our experiences in our social groups and cultures? Look for some very interesting examples of both.
- How stable are people's lay theories? Can we shift them to suit our goals or needs at the moment? Hint: They can shift in fascinating ways!

The next section addresses the consequences of lay theories about human psychological attributes or phenomena.

- What lay theory about willpower makes us want to push through and continue working rather than rest?
- What lay theories make us want to confront prejudice or injustice rather than throwing our hands up and moving on?
- You know how unusual thoughts can just pop into your head when you least expect them? Is this a good thing or a bad thing? Can we control them and, if so, should we try to?
- Are creative people born or made? Lay people and researchers (the experts) disagree. What do the experts think?

The third section highlights the consequences of lay theories about the metaphysical or supernatural:

- Do you think the question of mind–body dualism is relevant only to philosophers? If so, take a look at how believing in mind–body dualism can foster unhealthy eating.
- When are people most likely to produce magical explanations for something that happened? Hint: It is not about miracles.
- How do our religious beliefs shape so many aspects of our lives, including our self-regulation, risk-taking, and relationships?

And the final section deals with the consequences of lay theories about mental and physical health or illness.

- Is thinking about cancer as “a war with an enemy” good or bad for you?
- When therapists think of a mental disorder (such as depression) as biological, will they become more or less sympathetic to their patients?
- Is obesity something that is written in our genes or is it something we have personal control over? What are the advantages of these different points of view?

See what I mean? You will find yourself delving into one chapter after another, learning fascinating and valuable things about people’s deepest beliefs and the impact of these beliefs on all aspects of their lives. You will find yourself relating the chapters to each other and asking new questions. And you may well be tempted to try your hand at research on implicit theories yourself.

Carol S. Dweck
Department of Psychology, Stanford University, Stanford, USA

Contents

Foreward

Carol S. Dweck

Part I The Origins and Nature of Lay Theories

The Origins of Lay Theories: The Case of Essentialist Beliefs 3

Nick Haslam

The Motivated Fluidity of Lay Theories of Change 17

Anne E. Wilson and Jaslyn A. English

Part II Explorations in Lay Theories About Human Psychological Attributes or Phenomena

Lay Theories of Self-control 47

Veronika Job and Gregory M. Walton

What Are People’s Lay Theories About Mind Wandering and How Do Those Beliefs Affect Them? 71

Claire M. Zedelius and Jonathan W. Schooler

Lay Theories of Creativity 95

Simone M. Ritter and Eric F. Rietzschel

Mindsets About Malleability and Intergroup Relations 127

Aneeta Rattan and Oriane Georgeac

Effects of Lay Beliefs on the Justice Motive 157

Michèlle Bal and Kees van den Bos

Part III Insights into Lay Theories About the Metaphysical or Supernatural	
Antecedents, Manifestations, and Consequences of Belief in Mind–Body Dualism	181
Matthias Forstmann and Pascal Burgmer	
Lay Theories of the Mind/Brain Relationship and the Allure of Neuroscience	207
Diego Fernandez-Duque	
Causes and Consequences of the Belief in Free Will	229
Davide Rigoni, Axel Cleeremans and Marcel Brass	
Religion and Its Cultural Evolutionary By-products	243
Kristin Laurin	
From the Impossible to the Improbable: A Probabilistic Account of Magical Beliefs and Practices Across Development and Cultures	265
Martin Fortier and Sunae Kim	
Part IV Investigation of Lay Theories About Mental and Physical Health	
Mindsets of Body Weight	319
Jeni L. Burnette, Crystal L. Hoyt and Kasey Orvidas	
Lay Theories and Metaphors of Health and Illness	341
David J. Hauser, Randolph M. Nesse and Norbert Schwarz	
How Lay Theories Influence Our Mental Health	355
Adrian Furnham	
Index	375

Contributors

Michèle Bal Department of Interdisciplinary Social Science, Utrecht University, Utrecht, The Netherlands

Kees van den Bos Department of Psychology and School of Law, Utrecht University, Utrecht, The Netherlands

Marcel Brass Department of Experimental Psychology, Ghent University, Ghent, Belgium

Pascal Burgmer Department of Psychology, University of Cologne, Cologne, Germany

Jeni L. Burnette Department of Psychology, North Carolina State University, Raleigh, NC, USA

Axel Cleeremans Consciousness, Cognition and Computation Group, Brussels, Belgium; Center for Research in Cognition and Neurosciences, Brussels, Belgium; ULB Neurosciences Institute, Université Libre de Bruxelles, Brussels, Belgium

Jaslyn A. English Psychology Department, Wilfrid Laurier University, Waterloo, ON, Canada

Diego Fernandez-Duque Psychology Department, Villanova University, Villanova, PA, USA

Matthias Forstmann Department of Psychology, Yale University, New Haven, CT, USA

Martin Fortier Ecole Normale Supérieure, Ecoles des Hautes Etudes en Sciences Sociales, Institut Jean Nicod, PSL Research University, Paris, France

Adrian Furnham Department of Clinical, Educational and Health Psychology, University College London, London, UK

Oriane Georgeac Organisational Behaviour, London Business School, London, UK

Nick Haslam School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia

David J. Hauser Department of Psychology, University of Michigan, Ann Arbor, MI, USA

Crystal L. Hoyt Jepson School of Leadership Studies, University of Richmond, Virginia, USA

Veronika Job Department of Psychology, University of Zurich, Zurich, Switzerland

Sunae Kim Institute of Psychology, Department of Developmental and Clinical Child Psychology, Eötvös Loránd University, Budapest, Hungary

Kristin Laurin Department of Psychology, University of British Columbia, Vancouver, BC, Canada

Randolph M. Nesse Center of Evolution and Medicine, Arizona State University, Tempe, AZ, USA

Kasey Orvidas Department of Psychology, North Carolina State University, Raleigh, NC, USA

Aneeta Rattan Organisational Behaviour, London Business School, London, UK

Eric F. Rietzschel Department of Psychology, University of Groningen, Groningen, The Netherlands

Davide Rigoni Department of Experimental Psychology, Ghent University, Ghent, Belgium

Simone M. Ritter Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

Jonathan W. Schooler Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, USA

Norbert Schwarz Department of Psychology, Marshall School of Business, Dornsife Mind & Society Center, University of Southern California, Los Angeles, CA, USA

Gregory M. Walton Department of Psychology, Stanford University, Stanford, CA, USA

Anne E. Wilson Psychology Department, Wilfrid Laurier University, Waterloo, ON, Canada

Claire M. Zedelius Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, USA

Part I
The Origins and Nature of Lay Theories

The Origins of Lay Theories: The Case of Essentialist Beliefs

Nick Haslam

According to a well-known saying, “we see things not as they are, but as we are.” This aphorism distils two basic psychological truths: perception constructs our sense of reality, and people perceive the world in different ways. If anything, the saying does not go far enough. We not only construct our mental reality and do so in different ways, but we also alter our physical and social reality by acting on our perceptions. The fertile concept of lay theories shows how this world-making process takes place. People hold different beliefs about such things as intelligence and personality, they make sense of their experiences differently based on those beliefs, and as a consequence they think, feel, and behave differently, in ways that may change their life outcomes.

The origins of my opening quote are uncertain. It is sometimes attributed to the Talmud, sometimes to the novelist Anaïs Nin, and sometimes to the philosopher Immanuel Kant. The same uncertainty surrounds the origins of lay theories themselves. Most researchers approach them as already existing beliefs whose correlates and effects can be studied in the present. However, they are also the products of particular psychological structures and mechanisms, and outputs of particular developmental processes. In the present chapter, I attempt to answer the question of where lay theories come from. My focus is on the origin of those theories that involve psychological essentialism, the belief that something has an inner essence or nature that determines its identity and its outward features. I argue that the origins of these lay theories must be approached from several different directions. In brief, essentialist lay theories originate in fundamental cognitive tendencies of the human mind, in particular developmental experiences and inputs, in particular cultural settings, and in particular social arrangements.

N. Haslam (✉)

School of Psychological Sciences, University of Melbourne,
Melbourne, VIC 3010, Australia
e-mail: nhaslam@unimelb.edu.au

The chapter begins with a brief review of work on psychological essentialism and how the concept of lay theories relates to essentialist thinking. It then proceeds to discuss the origins of essentialist lay theories from these four perspectives: cognitive, developmental, cultural, and social. I conclude by arguing that a full understanding of the origins of lay theories requires that all of these perspectives are taken into account. Like the saying, lay theories have multiple authors.

Psychological Essentialism

Research on psychological essentialism arose first within cognitive and developmental psychology. Cognitive psychologists Medin and Ortony (1989) argued that although essentialism is generally thought by philosophers to be bad metaphysics—categories rarely if ever have underlying essences—essentialist intuitions are held by many laypeople. ‘Psychological essentialism’ refers to these intuitions. Everyday people tend to believe that what makes cats cats is some sort of feline essence that is shared by all members of the species. This deep-seated hidden essence underlies the observable characteristics of cats—their appearance and their behavior—and it determines their identity. People may have no concrete idea of what that essence might be, but that are confident that it exists and that an appropriate expert knows the answer.

Developmental psychologists such as Gelman (2003) and Keil (1994) demonstrated that young children possess essentialist intuitions about living kinds. Keil showed that children believe that a member of one species retains its species identity even if its appearance is transformed so that it resembles another species. In contrast, human artifacts such as items of furniture were not judged to retain their identity when they were similarly transformed. The key difference at play is that children believe that living kinds have some sort of inner essence or nature that endures despite outward alterations, whereas human artifacts do not. Essences are therefore ways of explaining the immutability of identity: the essential nature of a thing endures despite changes to its outward appearance, and because identity rests on this essence the thing itself is unaltered.

Social psychologists came somewhat late to the study of psychological essentialism. Cognitive and developmental psychologists had emphasized ‘natural kind’ concepts such as chemical elements or biological species as the focus of essentialist intuitions, rather than kinds of person or human attributes. However, important theoretical work by Rothbart and Taylor (1992) argued that people often hold essentialist intuitions about some human groups, and especially those associated with differences of appearance such as race and gender. These intuitions, Rothbart and Taylor proposed, amount to a failure by everyday folk to recognize that human groups are artifacts rather than natural kinds. To believe that racial categories are grounded in hidden essences is to mistakenly view these categories as timeless and species-like, when they are in fact contingent products of history and culture.

By Rothbart and Taylor's reckoning, essentialist beliefs about human groups are factually wrong and socially destructive, as they lead people to infer deep and unalterable differences between people based on superficial differences in their appearance.

Rothbart and Taylor (1992) inspired a lively empirical literature on essentialist beliefs about human groups and attributes. That literature is too large and complex to summarize here but a few key findings can be extracted. First, essentialist beliefs about social groups are multifaceted. They involve beliefs that the group has inherent characteristics, that it is in some sense 'natural' or biologically based, that it is highly informative about its members, that all members of the group are fundamentally alike, and beliefs that membership in the group is discrete (either/or), immutable, and historically invariant (Haslam, Rothschild, & Ernst, 2000). Second, the implications of essentialist beliefs are generally negative: people who hold more essentialist beliefs about a group tend to be more prejudiced toward that group, more reluctant to cross group boundaries (e.g., interact with people of other racial backgrounds), more resistant to egalitarian intergroup relations, more prone to endorse group stereotypes, and even less creative (e.g., Bastian & Haslam, 2006; Haslam, Rothschild, & Ernst, 2002; Keller, 2005; Tadmor, Chao, Hong, & Polzer, 2013; Williams & Eberhardt, 2008). Researchers have now explored the structure and implications of essentialist beliefs about numerous social groups, including those based on gender, race, sexuality, and mental disorder, as well as beliefs about some human attributes, most notably personality (Haslam, Bastian, & Bissett, 2004). The fundamental difference between these lines of work is that groups are conceptualized as noun classes, whereas attributes are conceptualized primarily as properties that vary by degree and do not constitute categories.

It is here that research on psychological essentialism makes contact with research on other lay theories. Although the concept of lay theories is very broad, as this volume attests, my focus in this chapter is on the specific conception of lay theories developed by Carol Dweck and colleagues (e.g., Dweck, 1999; Dweck & Leggett, 1988), and how it relates to essentialist lay theories. Dweck and her fellow researchers have identified a crucial way in which people's beliefs about human attributes such as intelligence and personality differ. 'Incremental theorists' believe that these attributes are malleable, dynamic, and always in flux or in process. 'Entity theorists', in contrast, believe that these attributes are fixed, static, and entity-like products. Fundamentally, the difference between holders of these two kinds of theory is that some believe that people can change and some do not. Because immutability is a key element of essentialist thinking, it is reasonable to ask whether Dweck's lay theories can be viewed through the conceptual lens of psychological essentialism and whether holding an entity theory of a human attribute is tantamount to holding essentialist beliefs about it.

In past work, my colleagues and I (Haslam, Bastian, Bain, & Kashima, 2006) have explored these questions and argued in the affirmative. We have proposed that Dweckian lay theories pick out one crucial element (immutability) in an interlinked set of essentialist beliefs. Believing that personality is fixed can be taken as an entity

theory of personality, or it can be taken as simply one component of a set of essentialist beliefs that personality is fixed and also biologically based, inherent, informative, discrete, and so on. There is evidence that these beliefs about personality do in fact covary in a coherent fashion (Haslam et al., 2004), and also that holding an entity theory about personality tends to correlate with believing that this continuity is caused by inhering qualities of the person (Haslam, Bastian, Fox, & Whelan, 2007). Similarly, the standard measure of entity theories about personality correlates with measures of other essentialist beliefs (Bastian & Haslam, 2006). Dweck's concept of entity theories also aligns with the broader concept of psychological essentialism in another respect. Her colleagues' work consistently demonstrates that entity theories have damaging implications, leading people to avoid academic challenges, to get anxious in assessment contexts, to endorse social stereotypes, and to form premature impressions of other people (e.g., Levy, Stroessner, & Dweck, 1998). These negative ramifications of entity theories are entirely compatible with the almost uniformly negative implications of essentialist beliefs in the social domain.

For the remainder of this chapter, I will consider entity theories in the Dweckian sense as particular forms of essentialist belief. I fully recognize that Dweck's framework does not exhaust the range of lay theories, and that her incremental theories are non-essentialist. Viewed in the manner, I propose, an entity theory is a minimal essentialism-related belief that singles out one important component of essentialism—immutability—and does not specify the supposed cause of the inability to change. An entity theory, as a belief that some human attribute is fixed, is simply one element of a multifaceted set of essentialist beliefs about that attribute. This theory will usually be accompanied by a belief that the immutability of the attribute is caused by an underlying and inhering essence of some sort. An incremental theory, on the other hand, represents one component of a set of nonessentialist beliefs about human attributes.

Understood in this way, the question of the origins of lay theories becomes the question of the origins of essentialist thinking. I propose that these origins must be approached from four distinct vantage points. First, essentialist thinking has particular origins as a mode of cognition. These *cognitive* foundations of psychological essentialism represent the distal origins of lay theories. Second, essentialist thinking about particular human groups and attributes arises in part in response to particular *developmental* experiences, such as the ways in which particular groups are described in language and the communications that parents have with their children about everything from ethnic diversity to their academic performance. In the absence of these experiences, the essentialist mode of cognition is unlikely to emerge. Thus, these developmental influences represent more proximal origins of lay theories. A third origin of essentialist thinking can be found in people's ambient culture, which supplies the idioms through which we understand essences, whether these be biological, spiritual, or something else. Certain ways of conceptualizing essences are particularly available in particular *cultural settings*, and these idioms shape the content of lay theories. Fourth and finally, to understand the origins of lay theories as forms of essentialist thinking we need to attend to the prevailing

social arrangements. Essentialist thinking about human groups and attributes is promoted by particular on-the-ground social realities, and these realities must be recognized in any complete account of the origins of lay theories.

Cognitive Foundations

Theories of the origins or psychological essentialism, either as a general phenomenon or specifically as they concern beliefs about human attributes and groups, propose that they are grounded in more fundamental cognitive tendencies. Three main accounts of these cognitive foundations of essentialist lay theories have been proposed.

The most prominent account of essentialist thinking as it applies to social groups proposes that it is grounded in folk biology, an evolved way of thinking about living kinds such as biological species. The natural world of plants and animals is composed of discrete biological categories or ‘natural kinds’. Members of one species do not suddenly transform into members of another—species membership is an unalterable part of their identity—and species themselves appear to be timeless. According to theorists such as Atran (1990), people have folk biological intuitions that represent nonhuman species in an essentialist fashion. By this account, when people hold essentialist lay theories about human groups, such as believing races to be timeless biologically based categories, they are extending folk biological ways of thinking into the social domain. That is, essentialist theories of human groups represent those groups as if they were distinct biological species.

One example of this view comes from the work of Gil-White (2001), who conducted ethnographic fieldwork on essentialist thinking about ethnicity in Mongolia. Gil-White presented participants from two ethnic groups who lived in the region with vignettes in which a baby born to biological parents of one group is raised by parents from the other group. Participants were asked whether the baby would grow up to have attributes of each group, pitting nature and nurture against one another in a hypothetical adoption study. Finding that study participants commonly expected that children would grow up to embody the ethnic attributes of their biological parents despite not being reared by them, Gil-White inferred that they held essentialist intuitions about ethnicity. Like the ugly duckling, a cygnet raised by ducks whose true swan nature is revealed as it matures, Mongol-born babies will come to display typically Mongol behavioral traits even when they have been reared by Kazakh parents. Gil-White theorized that these intuitions make good sense in the light of folk biology. Just as members of biological species mate with one another (endogamy) and always give birth to members of their species (descent), members of ethnic groups also tend to marry within their group and transmit their ethnicity to their children. Essentialism is a not unreasonable interpretation of these arrangements. Gil-White went further to argue that essentialist thinking about groups may have adaptive benefits, as it discourages potentially costly and hard to coordinate intergroup interactions.

A second account of the origins of essentialist thinking challenges the view that it is rooted in folk biology. Instead, this account maintains that essentialism is one of several ‘modes of construal’ that are available for making sense of the world—alongside teleological, mechanistic, and intentional modes, for example (Keil, 1994)—and it can be applied in a variety of domains rather than being intrinsic to one (i.e., biology). If people understand differences between people in an essentialist fashion, viewing them as deeply rooted and unchanging, they may be doing so for reasons other than a perceived analogy with interspecies differences.

Hirschfeld (1996), for example, argues that racial essentialism does not arise out of an analogical transfer of folk-biological intuitions about discrete, essence-based species into the realm of phenotypic variations among humans. For a start, these variations would have to be classified into types before the analogical transfer takes place, and in addition, these phenotypic variations do not correspond to discrete categories as obviously as biological species do. Rather than folk-biological essentialism being imported into the social domain, essentialist thinking is simply an available way of construing phenomena, and people are in some way prepared to encounter ontologically distinct types in the social domain.

A third account of the cognitive foundations of psychological essentialism has been proposed more recently by Cimpian and Salomon (2014). They propose that an important precursor to psychological essentialism is a hitherto unappreciated cognitive process that they dub the “inherence heuristic.” This heuristic represents a tendency to appeal to the inherent or intrinsic features of an entity, such as its visible properties, when thinking about why it is the way it is. The heuristic is prior to essentialism because it need not involve any belief that the intrinsic features reveal an underlying essence. For example, the heuristic leads people to believe that there is something natural, right, or appropriate about things being as they are based on their intrinsic features (e.g., it is natural for people to have orange juice for breakfast because of the way it tastes). According to Cimpian and Salomon, people tend to over-rely on inherent perceptible features because they are more cognitively available and salient than extrinsic features, such as the entity’s history (e.g., how orange juice came to be a breakfast staple in some societies) or its relationships with other entities, which are in principle no less adequate as explanations.

The inherence heuristic is a broad tendency that leads people to give special weight to the properties that inhere in objects when explaining their actions. When it is applied in the social domain to explain the behavior of persons rather than things, it leads us to overestimate the degree to which behavior springs from internal characteristics of the actor rather than external aspects of the situation. The heuristic may therefore play a role in well-known social psychological effects such as the correspondence bias (Gilbert & Malone, 1995) at the level of attributions for individual behavior, and system justification (Jost & Banaji, 1994) at the level of beliefs about groups. Importantly, Cimpian and Salomon argue that the heuristic is a basis for essentialist thinking. According to this view, psychological essentialism is an elaboration of the tendency to explain people’s behavior in terms of their inherent qualities, in which those qualities are conceptualized as inner essences. In support of this claim, Salomon and Cimpian (2014) have shown that people who

rely more heavily on the inherence heuristic, assessed by endorsement of inherence-based intuitions (e.g., “There are good reasons why dollar bills are green”) are more likely to endorse essentialist beliefs about an assortment of human groups. In addition to this correlational evidence, they also showed that an experimental manipulation that diminished the use of the heuristic by priming extrinsic explanations for societal patterns reduced endorsement of essentialist beliefs. Thus, the inherence heuristic offers a plausible account of the cognitive origins of essentialist lay theories, and also of a cognitive process that is associated with individual differences in essentialist thinking.

Developmental Experiences

We have seen how folk biology, basic modes of construal, and the inherence heuristic provide different accounts of the origins of essentialist thinking. The three explanations point to alternative cognitive foundations on which psychological essentialism may be built. Lay theories which hold that a particular human group or attribute is immutable, like Dweck’s entity theory, are ultimately based on a tendency to view human variations as akin to biological species, on the deployment of a basic essentialist mode of construal, or on the tendency to explain phenomena in terms of the inherent properties of the people involved. These accounts clarify the distal origins of essentialist thinking in the abstract, but they fail to explain how some individuals come to hold essentialist beliefs more than others. To understand the more proximal sources of essentialist lay theories we must examine the developmental experiences that contribute to essentialist thinking.

Although some of the influences that promote the development of essentialist thinking may be nonverbal, most research on the subject has emphasized the role of language and language use. Three factors in particular have been identified as contributors to essentialist thinking among children. The first of these factors is the use of noun classes. Although the use of nouns does not entail essentialist beliefs about the categories to which they refer, there is evidence that nouns do lead people to infer that category membership is stable. Research by Gelman and Heyman (1999), for example, showed that when ostensibly the same information about a group is presented as a noun label rather than as an adjective or verbal phrase—describing its members as “carrot-eaters” rather than “people who eat carrots whenever they can,” for example—children draw different inferences about the group. In particular they infer that the group is coherent and stable over time. Nouns imply that the category refers to unchanging attributes and fixed identities. The use of nouns to refer to social groups and human attributes therefore probably evokes children’s capacities to essentialize categories.

A second influence on essentialist thinking is the use of generic statements such as “girls are kind.” Generics are consistent with essentialist category representations because they imply that categories are homogeneous and have inherent properties. Generic expressions are commonly used in all languages, but the extent to which

they are used varies widely between people and appears to have implications both for promoting essentialist thinking and for transmitting it from adults to children. Gelman, Ware, Kleinberg, Manczak, and Stilwell (2014) provided strong evidence for these claims. In a study of parents and their 2- to 4-year old children, they demonstrated that parents and children both showed consistent individual differences in their use of generics, and that parents' tendencies to produce generics correlated with those of their children. Moreover, parents who held more essentialist beliefs about traits tended to use more generics and to have children who did so, raising the possibility that the essentialist beliefs of adults may be transmitted to children via the production of generalizing statements.

This possibility receives support from the work of Marjorie Rhodes, Sarah-Jane Leslie, and Christina Tworek (2012), who showed experimentally that both 4-year olds and adults tended to develop essentialist beliefs about a novel social category ("Zarpies") that had been described using generic rather than specific language (i.e., "Zarpies are scared of ladybugs" versus "This Zarpie is scared of ladybugs"). In another experiment, these researchers established that inducing adults to hold essentialist beliefs about another novel category—telling them that Zarpies are "a distinct kind of people with many biological and cultural differences from other social groups" (p. 13528)—led them to use more than twice as many generics when talking about members of the category while showing a picture book about them to their children. As Gelman et al. (2014) argued, "some people seem to be more likely to think about the world as consisting of stable categories, and this way of thinking is reflected in their language use" (p. 936), which may in turn replicate essentialist thinking in the small consumers of that language. Parents may therefore create the proximal linguistic environment in which their children's distal tendency to think in terms of essences can flourish.

Communications about social categories can promote essentialist thinking by means other than generics. Gelman, Taylor, and Nguyen (2004) conducted a microanalytic analysis of parent-child talk about gender and found that even gender-egalitarian mothers may unwittingly promote essentialist thinking about gender in their young children. They do so not only by using generics but also by using gender labels and accentuating gender differences. Parent-child conversation may, therefore, spread essentialist thinking by conveying the sense that gender categories are internally homogeneous, stable, coherent, and informative.

Essentialist thinking about human attributes like intelligence and personality, rather than about categories such as race and gender, may have different conversational contributors. Several studies indicate that feedback given to children on their performance may promote beliefs that certain attributes are stable or fixed. Feedback that refers to static person attributes rather than dynamic processes—"you are good at that" rather than "you must have tried hard at that"—appears to encourage an entity theory of those attributes and the motivational frameworks associated with this theory. Mueller and Dweck (1998), for example, showed that praise for ability ("person praise") among fifth graders promoted a theory of intelligence as fixed and innate, and led to a lack of persistence following failure. Similar findings have been obtained among kindergarten children (Zentall &

Morris, 2010). There is even some evidence that children who received more “process praise” from their parents at ages 1–3, coded from recordings of spontaneous interactions in the home, held less fixed understandings of ability five years later (Gunderson et al., 2013).

In sum, there is strong evidence that developmental experiences exert a proximal influence on the development of essentialist thinking. These experiences, represented by exposure to particular forms of language use, promote essentialist beliefs about social categories and entity theories of human attributes. The common thread in these experiences is that they invoke a view of social groups and persons as fixed and consistent. Intriguingly, apparently innocuous and everyday uses of language may play a role in engendering essentialist thinking. Simply by using race labels, making nonsexist generalizations about gender, or offering generous but ability-focused praise we may be nourishing the child’s tendency to see a social world populated by deep divisions and stable hierarchies.

Cultural Settings

Deep-seated cognitive tendencies lay the foundation for essentialist lay theories. Exposure to certain linguistic expressions and forms of language use build on these foundations, leading people to hold essentialist theories about some phenomena more than others and leading some people to hold more essentialist theories than their peers. However, the belief that some sort of essence underlies a phenomenon can be expressed in many distinct ways depending on what that essence is understood to be. Essences are by their very nature unobservable, and the intuition that a hidden essence or nature lurks beneath the surface of observable phenomena often occurs in the absence of any clear sense of what that essence might be. It is for this reason that cognitive psychologists have referred to essentialist beliefs as invoking an “essence placeholder.” Because laypeople’s intuitions about the nature of hidden essences are often cloudy and minimal, they can be explicitly conceptualized in quite different ways. The supposed content of these hidden essences is therefore likely to be drawn from the repertoire of explanations that are salient in a particular culture at a particular time. Thus, culture supplies some of the idioms through which people make sense of the hidden nature of things.

Consider the case of lay theories of race or ethnicity, for example. The belief that humans belong to an assortment of fundamentally different racial or ethnic types has probably always been widespread. People observe superficial but correlated variations in appearance, customs, and languages and infer that beneath the surface human groups have deep differences in kind. However, the basis of these underlying differences is mysterious and must be given an explanation that makes sense within a cultural context. Among people who believe that the world is populated by unseen spirits, the belief that different human groups have different spiritual essences—such as animal spirits that reveal their descent from different animal

lineages—is likely to be a culturally satisfying explanation. In cultures with a more materialistic mindset, racial and ethnic essences might be understood through the metaphor of blood: different groups have different blood and mixing these different kinds of blood is a cause for concern. In a genomic age, genes may become a preferred idiom for understanding group essences.

These disparate cultural idioms give different content to essentialist intuitions, and they therefore represent different lay theories with potentially diverging consequences. Understanding the hidden nature of a group in terms of its spirit, its blood, or its genetic make-up has different implications for how members of the group acquire or may lose their essence. It may have different implications for whether the group can ever merge with other groups, and for whether it is a meaningful part of a divine plan or merely a product of the blind forces of evolution. However, despite these differences, spiritual, sanguinary, and genetic essences are alike in other ways. They all provide a means of explaining how unobservable causal factors can account for deep and enduring differences between people.

Genetic lay theories are an interesting case in point. Such theories are popular in modern industrialized societies owing to the cultural salience and scientific respectability of genetics and the declining belief in spiritual explanations in a materialist age. Genetic lay theories are of course historically recent given the relatively recent discovery and characterization of DNA. However, although such theories have the trappings of modern science, they are often employed in crudely prescientific ways, functioning in the same essentialist way as ideas of “blood” (Dar-Nimrod & Heine, 2011). Genes and DNA commonly serve as culturally accepted idioms for expressing simple essentialist intuitions. Thus, genes are often seen as deterministic causes that have large, binary, and unchangeable effects, although none of these beliefs accurately captures the truth of most genetic influences on human behavior. These influences typically involve multiple genes of small effect that give rise to continuously distributed phenotypes and that are modulated in complex, nondeterministic ways by the environment.

Genetic essentialism (Dar-Nimrod & Heine, 2011) is now a prevalent form of lay theory of an assortment of human differences. Genetic essentialist theories are especially common explanations of variations in appearance such as race, ethnicity, gender, and obesity, but they are also prevalent in people’s thinking about less visible differences, such as mental disorders (Haslam, 2011; Kvaale, Haslam, & Gottdiener, 2013), sexual orientation (e.g., Haslam & Levy, 2006), and personality (Haslam et al., 2004). The content of these lay theories gives them the lustre of science, but in many respects their implications are no more enlightened than archaic ideas of blood. For example, people who hold more genetically essentialist beliefs generally tend to display higher levels of ethnic prejudice (Keller, 2005) and more stigmatizing attitudes toward the mentally ill (Kvaale et al., 2013). Genetic lay theories also have problematic implications for people who experience health problems such as obesity, indicating that they are difficult to change and that personal efforts to do so are likely to fail (see Burnette, Hoyt, & Orvidas, 2017;

Dar-Nimrod, Cheung, Ruby, & Heine, 2014). Indeed, in some respects genes and DNA may provide a more destructive basis for essentialist lay theories than alternative, older ways of construing essences. Spirits may be appeased or ritually exorcized and blood may be cleansed and purified, but genes are understood to be unalterable. Genetic factors are also understood as less malleable than neural factors, another equally modern explanation that can be seen as essence-like (“neuroessentialism”; Haslam, 2011). Thus, genes may be an especially potent foundation for what Dweck and colleagues refer to as entity theories of human attributes, with all the well-documented damaging implications that are associated with such theories.

Regardless of the pros and cons of particular ways of representing essences, my key point is that essentialist lay theories partly originate in these idioms. Our cultures provide particular idioms for explaining the underlying causes of phenomena and these idioms shape how essentialist theories are expressed and how they influence behavior, social interaction, and intergroup relations.

Social Arrangements

Culture supplies the contents of essentialist lay theories by providing preferred ways of explaining what the underlying essences might be. However, lay theories of human attributes and groups are also likely to be constrained by existing social arrangements. Essentialist theories are more likely to take root in some societies than in others, and some social arrangements are especially likely to support these theories. In addition, people may be motivated to hold essentialist lay theories out of a desire to maintain existing social arrangements.

There has been little research on whether particular social arrangements in some way encourage or support essentialist thinking about social groups and human attributes. However, there are several grounds for believing that this is a plausible claim. In particular, because essentialist thinking envisions a social world composed of fixed entities, static hierarchies, and deep intergroup divides, we might expect that societies in which social divisions are relatively stark and durable might promote such thinking. For example, a belief in racial essences may be more likely to endure in a context of lasting residential and occupational segregation, and limited interaction and intermarriage (cf. Gil-White, 2001, on the principle of descent in ethnic essentialism). As much as this social segregation is in fact an artifact of history rather than a fact of nature, it should be relatively easy to construe as the natural manifestation of distinct groups with fundamentally different essences. It should be more difficult to develop and sustain a belief in racial essences in social contexts where racial diversity is less dichotomous, more fluid, and less of a barrier to intermingling, as in Brazil. In a similar fashion, gender essentialism should be more likely to take hold—or be harder to dislodge—in social contexts where sex roles are relatively rigid and distinct. We might imagine a mutually reinforcing

tendency for segregated social arrangements to breed essentialist beliefs about the basis of those arrangements, and for endorsement of those beliefs to entrench those arrangements. For instance, women who hold more essentialist beliefs about gender are more likely to personally endorse stereotypically feminine traits (Coleman & Hong, 2008), and men who essentialize gender are less likely to provide direct care for their young children (Gaunt, 2006).

The reality of social arrangements may promote or inhibit essentialist thinking about groups and attributes, but people's positions within these arrangements may also play a role. There is some evidence that essentialist thinking sometimes represents motivated social cognition. Sometimes the motivation is primarily epistemic (e.g., driven by need for closure: Roets & Van Hiel, 2011) or self-serving, as demonstrated by evidence that people alter their theories about personal attributes in self-enhancing or self-protective ways (Leith et al., 2014; Steimer & Mata, 2016). However, sometimes the relevant motivation is social, driven by concerns about maintaining an advantageous position for one's group within existing arrangements. As system justification theory argues (Jost & Banaji, 1994), people are driven to rationalize the status quo by perceiving it as good, stable, and in the nature of things. Thus, men appear to be more likely to endorse genetic essentialism regarding gender than women (Keller, 2005), high caste Indians are more likely to agree with essentialist understandings of caste (Mahalingam, 2007), and higher social class Americans are more likely to endorse essentialist beliefs about class, an effect partially explained by their greater belief in a just world (Kraus & Keltner, 2013). Although the research has not been conducted, we might expect similar findings if essentialist beliefs about human attributes—that is, entity theories—were examined. People with higher measured intelligence may be more likely to believe that intelligence is stable and innate, for example, motivated by the desire to believe that their position high on the intellectual hierarchy is secure. Assessing the role of motivated essentialism in lay theories of human attributes would be an intriguing direction for future research.

Conclusions

My argument in this chapter is that lay theories which involve beliefs in the stability of social groups and human attributes can be usefully understood through the lens of psychological essentialism. That lens helps us to answer the question of where these theories come from. The origins of lay theories are complex and a full account of these origins has to approach the answer from several distinct angles. Those angles require us to address both distal and proximal factors and to appreciate the role of a tangled web of cognitive, linguistic, developmental, social, motivational, and cultural influences. The price of this inquiry into the origins of lay theories is complexity, but the benefit is a comprehensive understanding that connects lay theories research to a rich assortment of alternative research traditions.

References

- Atran, S. (1990). *Cognitive foundations of natural history*. New York, NY: Cambridge University Press.
- Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology, 42*, 228–235.
- Burnette, J. L., Hoyt, C. L., & Orvidas, K. (2017). Mindsets of body-weight. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Cimpian, A., & Salomon, E. (2014). The inherence heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences, 37*, 461–480.
- Coleman, J. M., & Hong, Y. (2008). Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity, 7*, 34–53.
- Dar-Nimrod, I., Cheung, B. Y., Ruby, M. B., & Heine, S. J. (2014). Can merely learning about obesity genes change eating behavior? *Appetite, 81*, 269–276.
- Dar-Nimrod, I., & Heine, S. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137*, 800–818.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia: Psychology Press.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.
- Gaunt, R. (2006). Biological essentialism, gender ideologies, and role attitudes: What determines parents' involvement in child care. *Sex Roles, 55*, 523–533.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.
- Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science, 10*, 489–493.
- Gelman, S. A., Taylor, M. G., & Nguyen, S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development, 69*(1, Serial No. 275).
- Gelman, S. A., Ware, E. A., Kleinberg, F., Manczak, E. M., & Stilwell, S. M. (2014). Individual differences in children's and parents' generic language. *Child Development, 85*, 924–940.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117*, 21–38.
- Gil-White, F. J. (2001). Are ethnic groups biological “species” to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology, 42*, 515–553.
- Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development, 84*, 1526–1541.
- Haslam, N. (2011). Genetic essentialism, neuroessentialism, and stigma: Comment on Dar-Nimrod & Heine (2011). *Psychological Bulletin, 137*, 819–824.
- Haslam, N., Bastian, B., Bain, P. G., & Kashima, Y. (2006). Psychological essentialism, implicit theories, and intergroup relations. *Group Processes & Intergroup Relations, 9*, 63–76.
- Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin, 30*, 1661–1673.
- Haslam, N., Bastian, B., Fox, C., & Whelan, J. (2007). Beliefs about personality change and continuity. *Personality and Individual Differences, 42*, 1621–1631.
- Haslam, N., & Levy, S. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin, 32*, 471–485.
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113–127.

- Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology, 41*, 87–100.
- Hirschfeld, L. A. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*, 1–27.
- Keil, F. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind* (pp. 234–254). New York, NY: Cambridge University Press.
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*, 686–702.
- Kraus, M. W., & Keltner, D. (2013). Social class rank, essentialism, and punitive judgment. *Journal of Personality and Social Psychology, 105*, 247–261.
- Kvaale, E., Haslam, N., & Gottdiener, W. (2013). The 'side-effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review, 33*, 782–794.
- Leith, S. A., Ward, C. L., Giacomin, M., Landau, E. S., Ehrlinger, J., & Wilson, A. E. (2014). Changing theories of change: Strategic shifting in implicit theory endorsement. *Journal of Personality and Social Psychology, 107*, 597–620.
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology, 74*, 1421–1436.
- Mahalingam, R. (2007). Essentialism, power, and the representation of social categories: A folk sociology perspective. *Human Development, 50*, 300–319.
- Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge, UK: Cambridge University Press.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*, 33–52.
- Roets, A., & Van Hiel, A. (2011). The role of need for closure in essentialist entitativity beliefs and prejudice: An epistemic needs approach to racial categorization. *British Journal of Social Psychology, 50*, 52–73.
- Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences, 109*, 13526–13531.
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 11–36). Thousand Oaks, CA: Sage.
- Salomon, E., & Cimpian, A. (2014). The inherence heuristic as a source of essentialist thought. *Personality and Social Psychology Bulletin, 40*, 1297–1315.
- Steimer, A., & Mata, A. (2016). Motivated implicit theories of personality my weaknesses will go away, but my strengths are here to stay. *Personality and Social Psychology Bulletin, 42*, 415–429.
- Tadmor, C. T., Chao, M. M., Hong, Y. Y., & Polzer, J. T. (2013). Not just for stereotyping anymore: Racial essentialism reduces domain-general creativity. *Psychological Science, 24*, 99–105.
- Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology, 84*, 1033–1047.
- Zentall, S. R., & Morris, B. J. (2010). "Good job, you're so smart": The effects of inconsistency of praise type on young children's motivation. *Journal of Experimental Child Psychology, 107*, 155–163.

The Motivated Fluidity of Lay Theories of Change

Anne E. Wilson and Jaslyn A. English

People encounter an almost overwhelming quantity of information about human behavior and the social world everyday. Despite this information overload, humans are markedly adept at finding signal in the noise, interpreting the inputs in their complex environments in a way that both simplifies and makes meaning. People use a variety of shortcuts or heuristics to make sense of these stimuli (Tversky & Kahneman, 1974); they also access a range of lay theories about the nature of humans, natural processes, and how the world works. Lay theories are sometimes referred to as naïve or folk theories acknowledging humans as naïve scientists attempting to make sense of a complex world (Heider, 1958; Kelly, 1955). These lay beliefs are also commonly called implicit theories in part due to the recognition that these beliefs often operate at an automatic rather than conscious level—people have assumptions, largely unexamined, about the world around them which guide their judgments, but which have rarely been articulated in careful detail or bolstered with rational argument. These implicit theories provide a lens through which people see the world and can shape their understanding of behavior, actions, and decisions in powerful ways.

This volume explores a wide range of these lay beliefs and articulates the many ways they can influence human thought, behavior and choice. To a large extent, these literatures tend to focus on how lay theories affect people's responses, either by examining individual differences in people's lay beliefs or by directly manipulating or creating the lay belief people hold. We will review only a small portion of this literature to paint a general picture of this approach. We focus primarily on one kind of implicit theory: people's beliefs about the fixed or malleable nature of

A.E. Wilson (✉) · J.A. English
Psychology Department, Wilfrid Laurier University,
75 University Avenue West, Waterloo, ON N2L 3C5, Canada
e-mail: awilson@wlu.ca

J.A. English
e-mail: engl3740@mylaurier.ca

human characteristics (Dweck, 2012), and extend our analysis to related theories regarding the mutability of groups and of social mobility. The central goal of this chapter, however, is to ask a slightly different question: are people's implicit theories chronic and stable over time, or do they shift in systematic ways? What leads people to adopt or endorse different lay theories at different times? We suggest that one set of important but unexplored factors pertains to people's current goals and identity needs. There may be times when people gravitate to one lay theory or another because a particular worldview will best help them to arrive at a particular, desired conclusion. We outline the emerging research examining, how people may shift their lay theories about the malleability of personal attributes in systematic ways when particular goals are activated. Although we focus primarily on these individual-level shifts in person lay theories, we will also consider how similar processes may play out in other domains in which competing lay theories of mutability have different implications for human behavior. Because, the literature on motivated adoption of implicit theories is limited, we will often make speculative connections that are not thoroughly tested. Our hope is to prompt additional research and theory in this area of study.

One of the implicit theories that have been studied extensively pertains to people's beliefs about the fundamentally fixed or malleable nature of human attributes (Dweck, 2012). Dweck describes entity theorists as believing that attributes are fixed and stable—people have a certain level of a given attribute or ability, and this level is relatively enduring. For instance, an entity theorist would strongly agree with the statement: *“Everyone is a certain kind of person, and there is not much that can be done to really change that.”* In contrast, incremental theorists are described as holding the conviction that people's attributes are inherently malleable with time and effort. For instance, an incremental theorist would strongly agree with the statement: *“People can change even their most basic qualities.”* These lay beliefs are often described in dichotomous terms (incremental and entity theorists), and for ease of communication we sometimes use these terms. Importantly though, people's actual views may fall anywhere on a continuum (typically a 6-point scale from Strongly Agree to Strongly Disagree). Further, although we sometimes will discuss incremental and entity theorists in general terms, in fact people hold different lay theories across domains (Dweck, Chiu, & Hong, 1995a). Someone may believe that morality is malleable but that intelligence is quite fixed. Some domains may be more closely associated than others, but in general, domain-specific lay theories will more accurately predict people's perceptions and choices (e.g., Chiu, Hong, & Dweck, 1997a; Ward & Wilson, 2015).

Why do these lay theories matter? There is considerable evidence that these theories guide person perception and stereotyping (e.g., Hong, Levy, & Chiu, 2001; Levy, Chiu, & Hong, 2006; Levy, Stroessner, & Dweck, 1998; Molden & Dweck, 2006), goal-pursuit and achievement (e.g., Burnette, Pollack, & Hoyt, 2010; Dweck & Leggett, 1988; Hong, Chiu, Dweck, Lin, & Wan, 1999), interpersonal relations and aggression (e.g., Kammrath & Dweck, 2006; Rattan & Georgeac, this volume; Yeager, Trzesniewski, & Dweck, 2013), and intergroup judgments (Jayaratne et al., 2006; Rattan, Savani, Naidu, & Dweck, 2012) among others. Specific phenomena

investigated vary across domains, but in general, implicit theories account for how people process complex social information. For instance, students who are incremental theorists are more likely to respond proactively to failure by seeking strategies for improvement, whereas entity theorists are less likely to select ameliorative strategies (Yeager & Dweck, 2012). Likewise, incremental theorists are more likely to approach conflict constructively (Kammrath & Dweck, 2006), and even attempt to change prejudiced attitudes (Rattan & Dweck, 2010) compared to entity theorists who believe such efforts would be futile if the interaction partner is unlikely to change.

It is less clear—though often hotly debated—which theory is more “correct” in reality. For instance, there is evidence for the hereditary nature of some temperaments, attitudes, and abilities (Harris, Vernon, Olson, & Jang, 1999). There is evidence that genetics may inhibit efforts at weight control (Bradfield et al., 2012; Burnette & Finkel, 2012). At the same time, evidence that people can change goes well beyond the inspirational stories of underdogs finding their way to success (Gladwell, 2013). Epigenetic research demonstrates the power of context and choice to determine how genetic factors are expressed (Sasaki, LeClair, West, & Kim, 2016), evidence shows how practice can change not only performance but the brain (Kelly & Garavan, 2005), and a great deal of social psychological research demonstrates the power of personal belief (Lou & Noels, 2016) and the situation to shape behavior over and above chronic dispositions (Reis, 2008). In short, there is plenty of evidence out there in the world for a reasonable person to draw on and conclude that attributes are quite malleable; there is also no shortage of evidence supporting the view that attributes are rather fixed. Beyond the world of research, Western cultural wisdom also contains mixed messages about the stable or dynamic nature of attributes. Proverbs have relegated thieves and leopards to a fate of perpetual sameness,¹ yet other wisdom purports that the “the only thing that is constant is change” (commonly attributed to Heraclitus). Although research evidence and folk wisdom often do not provide a singular answer to the question of which theory is more “correct,” evidence suggests that people tend to have an opinion. In surveys, about 80% of participants tend to report leaning more toward either an entity or an incremental viewpoint (Plaks, Levy & Dweck, 2009), with a relatively equal proportion endorsing each of the diverging viewpoints (Dweck, 2012).

We will make no claims about which theory tends to be more accurate (other than to say that “both” may often be the right answer). Regardless of accuracy, there is compelling evidence that the lay theory someone endorses about change can powerfully predict motivation, perception, and decisions. However, despite the large number of studies demonstrating meaningful consequences of implicit theories of malleability, it is unclear just how stable these theories are and what factors might influence a person’s dominant lay belief. First, how temporally stable are lay theories of change? We can answer this question in a few ways: by considering the

¹*Once a thief, always a thief; a leopard cannot change its spots.*

temporal stability of person theories, by considering how *malleable* lay theories are when presented with persuasive information, or by considering how they might change in response to specific experiences or contexts. We consider each in turn.

Temporal Stability of Implicit Theories

First, there is a general tendency to describe lay theories as relatively stable, chronic individual differences, implying that beliefs would remain quite consistent over time. There is some evidence supporting this contention: Dweck et al. (1995a) reported a test–retest reliability of 0.82 over 2 weeks on the 3-item Implicit Person Theory measure, and Levy et al. (1998) reported 0.82 over a week and 0.71 over 4 weeks for an 8-item measure. However, Poon and Koehler (2008) pointed out that the chronic, stable nature of implicit person theories is typically assumed rather than tested, and most often is measured either at the same time as dependent measures of interest or within a couple of weeks' time. Poon and Koehler examined stability over longer time periods, and found that the test–retest reliability declined considerably by 10 weeks out, down to 0.28. Further, they emphasized that contemporaneous measures of lay theories were strong predictors of relevant dependent variables. Specifically, in their research, implicit theories (measured contemporaneously) predicted intertrait inferences; entity theorists were more likely than incremental theorists to make inferences about a person's traits (e.g., *warm*) after learning that the person possessed a related trait (e.g., *sensitive*). However, when the lay theory measure was taken weeks earlier, it failed to consistently predict these same inferences, suggesting meaningful change in lay theories over time. Indeed, in a follow-up study they found that after 8 weeks, only about 60% of participants fell into the same entity or incremental theory category as they had at Time 1. As Poon and Koehler speculate, this temporal instability is worth noting when considering the chronic effects of implicit theories, that scores at any given moment are likely to involve "(a) one's chronic theory accessibility, as researchers have long assumed, but also (b) one's temporary theory accessibility triggered by naturally unfolding, idiosyncratic cues or experiences in everyday life" (Poon & Koehler, 2008, p. 975). Their conclusions emphasized the importance of their findings for research planning, as an earlier measure of lay theories might not adequately predict a later measure of outcomes. However, it led us to wonder about what kinds of day-to-day, idiosyncratic experiences may play a role in altering implicit theories. Were these fluctuations random, or systematic and explainable? In other research, Poon and Koehler (2006) describe implicit theories of change and stability as knowledge-activation frameworks: People likely possess knowledge of both lay theories, and may endorse different theories at different times as a result of the knowledge that has been become accessible in a given situation. They demonstrated that people readily shifted their theories after engaging in tasks designed to prompt them to search their memory for evidence of one theory or the other. For instance, people accessed malleability folk knowledge when asked to read a biography and

account for why the individual changed dramatically through life; they accessed entity knowledge when they explained a biography of someone who remained unchanged. Similarly, they activated knowledge consistent with different lay theories when asked to provide examples reflecting proverbs, such as “*You cannot teach an old dog new tricks*” or “*Experience is the best teacher.*” Further, the theory activated in the moment predicted subsequent unrelated trait judgments, demonstrating that people will make decisions and judgments on the basis of whatever theory is activated. They argue that stimuli akin to these kinds of experiences (person judgments, exposure to folk wisdom) are likely to occur in everyday life, accounting for some of the natural variation in people’s implicit theories over time. We concur, and speculate that people may vary in their implicit theory temporal stability in part depending on the contexts they find themselves in—it may be that some people find themselves (and select) circumstances that offer more evidence for stability on a day-to-day basis; others might encounter (or choose) environments illustrating change.

Experimental Malleability of Implicit Theories

There is no shortage of evidence that chronic implicit person theories can be changed. Indeed, the standard approach to establishing the causal effect of implicit theories is to (at least temporarily) experimentally manipulate the theory people hold. Most often, these theories are altered by presenting people with persuasive information, frequently in the form of a (bogus) research article that makes a strong case for either an entity or incremental understanding of a particular attribute (e.g., Chiu, Dweck, Tong, & Fu, 1997b; Hong et al., 1999; Levy, Stroessner, & Dweck, 1998; Molden, Plaks, & Dweck, 2006; Nussbaum & Dweck, 2008). This kind of overt persuasive argumentation bolstered by (ostensible) evidence appears to be quite effective at temporarily altering implicit theories and corresponding responses. In other research, researchers have attempted to alter these implicit theories longer term (Aronson, Fried, & Good, 2002; Blackwell, Trzesniewski, & Dweck, 2007). Because of their long-term focus, the researchers only attempted to shift people toward a more incremental view and not an entity one, given the preponderance of evidence suggests that an incremental theory offers more benefits. In longitudinal research, Yeager et al. (2013) focused on developing an incremental person theory in 9th grade students with the hypothesis that they would be less likely to attribute hostile intent behind ambiguous behaviors. Yeager et al. began by asking high school teachers to deliver a lecture about the malleability of the brain, bolstered by further scientific evidence and communication from peers 2 weeks later. Students were also asked to write notes to future classmates describing what they had learned. The control condition followed the same procedure but read about the malleability of athletic ability. Eight months later, they found that those in the experimental condition maintained an incremental perspective to a greater degree, and as expected, attributed less hostile intent than those participants in the control

condition. This provides some evidence that implicit theories may not just fluctuate but change directionally over time: in this case the shift was prompted by an initial set of persuasive communications but presumably was maintained by the way people came to actively process their environments (attending to and retrieving different information, interpreting incoming data through a particular lens, behaving in ways that would tend to confirm their existing theory). It is conceivable, then, that other real-world experiences may systematically prompt people to actively question, reassess, and possibly shift their lay theories in ways that would then tend to self-reinforce over time.

What Other Factors Affect Implicit Theories of Change and Stability?

We have evidence that implicit theories may not be especially temporally stable over time and that experimental manipulations can change them. Presumably, though, these implicit theories are shaped by other factors in people's environments as well. Understanding these mechanisms may give us clues to how these theories originate in the first place. We know that implicit theories can be subtly altered by the kind of feedback provided by parents and teachers (Gunderson et al., 2013; Mueller & Dweck, 1998): for instance, dispositional praise for achievements ("You're so smart!") may seem affirming, but may foster an entity theory in children who come to think of intelligence as a trait they possess. However, when these children encounter failure, they may then be more likely to attribute it to a lack of capacity. Children who are instead praised for the effort that went into achievement ("You must have worked very hard on that—good job!") are likely to tie success to hard work, and will be more inclined to view failure as a challenge to surmount with greater effort or different strategy (Dweck, Hong, & Chiu, 1993). These effects are meaningful especially given that parental praise in early childhood predicted children's motivational frameworks several years later (Gunderson et al., 2013)—and given that Mueller and Dweck (1998) report that a majority of parents believe it is important to praise ability following success to help children feel smart. This observation—that parents may offer counterproductive feedback because of an intuition that it may bolster self-esteem—offers an interesting insight that leads us to our next consideration. We know that self-image protection, maintenance, and enhancement processes can play a powerful role in how people actively process information, and that, in many instances, people are highly motivated to view themselves in a favorable light (Baumeister, 1998; Sedikides, 1993; Wood, Giordano-Beech, Taylor, Michela, & Gaus, 1994). Parents' intuition that praising children's innate abilities gives self-esteem a boost is far from baseless. Indeed, adults tend to fall into the same pattern of attribution when accounting for their own performance: research on the self-serving bias documents how people are much more likely to attribute their successes to dispositional factors (like their ability)

than their failures, which they are more likely to attribute to external causes (Campbell & Sedikides, 1999). At least one reason for this self-serving bias appears to be self-esteem maintenance (Shepperd, Malone, & Sweeny, 2008).

Motivated Fluidity of Lay Theories?

The parallel between lay theories of change and the self-serving bias suggests another mechanism by which implicit theories may shift over time. We know that people are active processors of the information available to them, and that often their processing is shaped by dominant motivations or goals. As theories of motivated reasoning suggest (e.g., Kunda, 1990; Pyszczynski & Greenberg, 1987; Taber & Lodge, 2006), people often begin the process of reasoning with a preferred conclusion already in sight. They also sometimes adopt different perspectives or principles to allow them to support the conclusion they most want to draw. We reasoned, then, that people may sometimes be motivated to shift their implicit theories to help them support their preferred conclusions. Imagine both Sarah and Alice got back grades on their math test. Sarah got an A, Alice got a D. For both women, math is relevant to their self-image. If both were then asked to consider whether intelligence is fixed or malleable, how might they each respond? Alice would have reason to gravitate toward an incremental theory, hoping that this grade does not seal her fate as a poor math student. Consistent with a knowledge-activation framework (Dweck, Chiu, & Hong, 1995b; Poon & Koehler, 2006) she may activate her existing knowledge around malleability, remembering instances where she has observed significant improvement in performance, and tell herself this is the kind of skill that can be mastered with hard work. Sarah, on the other hand, would not have this same motivation: she did very well on the test. She might congratulate herself by reminding herself how math ability is quite stable, so her performance likely heralds an enduring career of success. In this example, we suspect that Alice's motivation, after a threatening failure, may be stronger than Sarah's is after success, but both patterns would be largely consistent with a motive to protect or maintain self-esteem.

Although this kind of motivated fluidity seemed highly plausible in light of the existing literature, evidence for it seemed missing from our scan of the literature on Dweck's implicit theories. This prompted us to investigate these questions across a variety of contexts (Leith et al., 2014). We began by reasoning that people might be particularly likely to actively regulate their acceptance of these theories in response to particular types of situational goals. In particular, we thought that shifting lay theories would have its strongest intuitive appeal when people are faced with information about the self or others *over time* (Peetz & Wilson, 2008). That is, when people consider an individual's past attributes or behaviors, they must decide how it informs their present character. Likewise, people have to make judgments about whether past or current outcomes predict a person's future outcomes. In each of these cases, the lay theory one selects and applies to a given set of temporally

extended events can transform the conclusion. An entity theory suggests that past attributes reflect current character and, in turn, predicts similar future outcomes; an incremental theory presumes that people may have changes since the past point in time, and may likewise change in the future (Peetz & Wilson, 2014; Ross & Wilson, 2002). For instance, a past moral failure viewed in light of entity theory is likely to be seen as evidence of an enduring lack of trustworthiness, but through an incremental lens the same failure seems either irrelevant (since morality is changeable) or as information that can help foster growth. This argument is consistent with Kunda's (1990) and Pyszczynski & Greenberg's (1987) thinking about motivated reasoning, in which they argue that people cannot just believe whatever they want to believe in any moment, but rather they hold to an illusion of objectivity by engaging in a process of reasoning that involves the differential recruitment of knowledge, theories, and beliefs. More recent research supports this premise: people will appeal to different beliefs, convictions, and principles to support the conclusion they most prefer (e.g., Jost, Pelham, Sheldon, & Sullivan, 2003; Knowles & Ditto, 2012; Kunda, 1987; Schumaker & Slep, 2004; Skitka, Bauman, & Mullen, 2008; Tesser, 2001).

To test these ideas, we designed a series of experiments that fit the criteria we identified (Leith et al., 2014): Situations where people would be motivated to reach a particular directional goal (protecting the self, family, or important others) in which temporal information would be interpreted differently depending on the implicit theory: in other words, situations where being an entity theorist or an incremental theorist would lead to different conclusions on the basis of the same evidence. Next, we describe the evidence that endorsement of lay theories can be shaped by both self-image goals and other perception goals.

How Self-image Goals May Shape Implicit Theories of Stability and Change

We began by investigating contexts most directly connected to people's personal self-view, relying on the assumption that people would often be inclined to protect their self-view from threat (Leith et al., 2014). We created several situations in which people would have to face threatening information about the self: in two studies, we delivered a failure or success experience (feedback about poor/good performance on a test), and in another, we asked people to recall a personal memory representing a past social failure or success. In each of these cases, people were more likely to endorse an incremental theory about the nature of the attribute in question after encountering threatening rather than flattering information. That is, after getting a poor score on a test, people were more inclined to believe that ability was changeable with time and effort, whereas after success people were more willing to entertain the notion that these attributes were fixed and enduring.

In the first study we conducted (Leith et al., 2014, Study 1), we not only manipulated the outcome (success/failure following a bogus cognitive ability test), but also independently attempted to manipulate people's lay theories about the ability itself (describing the cognitive skill as highly malleable or quite fixed). This second manipulation mirrors the typical experimental interventions in the literature (e.g., Plaks & Stecher, 2007) and might be expected to alter people's dominant implicit theory on the basis of the information provided. We found evidence for both processes—a significant main effect revealed that the implicit theories manipulation did shift people's theory endorsement; people also endorsed incremental theories more strongly after failure than after success. Notably, an interaction also emerged, revealing that the implicit theory manipulation was effective in the success condition, but not in the failure condition. When unthreatened, people were willing to temporarily adopt whatever theory they learned about. However, after a threat (failure on a task reflecting their ability), people who were given persuasive information that the attribute was fixed were unconvinced by that argument—they endorsed a more incremental theory despite being presented evidence to the contrary. We speculate that this set of findings not only offers evidence that people may actively shift their dominant theory in situations when a particular perspective would help them to reach a desired conclusion, but also offers a caution to researchers seeking to alter lay theories—people may be more receptive to persuasive communication about the malleable or fixed nature of attributes when they do not have a motivation to be skeptical about that viewpoint (Taber & Lodge, 2006).

Although the finding across these studies suggests that people may actively shift toward a view of change or stability that supports their preferred interpretation of the evidence at hand, the fact that people support a more incremental view after failure than success does not in itself provide solid evidence that the effect is motivated. There are other possible reasons people might come to this conclusion—for instance, if an individual believes themselves to be highly competent in the threatened domain, then one piece of failure evidence might seem puzzling—they may conclude for more rational than motivated reasons that the attribute must be changeable given their fluctuating performance. We cannot rule out this process as among those that produce shifts in lay theories—indeed, it is quite likely. However, we did include more direct tests of the motivated nature of lay theory shifting. For instance, in one study (Leith et al., 2014, Study 3) we asked people to consider their own past failure or those of an acquaintance. People shifted their implicit theories only when considering personal outcomes, and not the outcomes of another individual whom they would have less inclination to protect. Of course, there are a host of differences between how we process information about self versus others—we have different amounts of information as well as different motivations. So, in a complementary approach (Leith et al., Study 2), we gave all participants self-relevant feedback (success or failure), but varied how meaningful the feedback seemed. Everyone completed a judgment task framed as a measure of “social perceptiveness.” The “thin-slice” person judgment task was engaging to participants, but entirely bogus. Participants were told that they performed exceedingly

well or poorly on the task. Then we attempted to alter the degree to which this feedback would be threatening. We described the test as a well-validated measure of consequential ability to one group of participants, and as a test that was still under development and unvalidated to another group. Those who had reason to believe the test was legitimate shifted their lay theories, and those who had an easy way to disregard the results as illegitimate did not. This evidence converges to suggest that at least one reason people's implicit theories may shift over time is due to the esteem-threatening experiences in their day-to-day lives that can be better incorporated into a positive self-view by shifting to an incremental lay theory.

The findings from Leith et al. (2014) were further supported by additional emerging research. Steimer and Mata (2016) asked people to list their strengths and weaknesses and to rate how likely those qualities were to change. Participants in their study professed a belief that only their own weaknesses were likely to change, but their own strengths were stable. This suggests that people can potentially hold both implicit theories virtually simultaneously, and simply view them as applying to different dimensions of identity. Is this perception motivated by self-goals, or do people hold a general theory that there are forces maintaining people's strengths and encouraging change on weaknesses? Although people shifted their beliefs about malleability when it came to *personal* weaknesses, participants viewed both the strengths and weaknesses of *other* individuals as relatively stable. This entity theory of others held even when the participant was told that the other person was motivated to change their weaknesses.²

Steimer and Mata (2016)'s findings are generally consistent with earlier research demonstrating how gifted students think about the malleability of their academic skills. Ziegler and Stoeger (2010) report that very successful students held both theories of change concurrently: successes were viewed through an entity lens, whereas failures and ability deficits elicited an incremental viewpoint. The authors interpreted these findings in terms of domain-specificity (success and failure as different domains) even though in many cases both were held for the same skill domain (e.g., math). Their findings are also consistent with a motivated fluidity account. To the extent that this fluidity is a particular feature of successful individuals, it also suggests that this flexibility in adopting various lay theories may serve more than self-esteem needs—it is possible that it also provides an adaptive

²However, it is also possible that people do not even hold the same implicit theories about self and others, although the general person scale seems able to predict both personal and other judgments. Some recent research developed a self-theories version of the implicit theories scale based on the recognition that people might have one belief about how malleable intelligence is in general, and a different view of their own personal intelligence. On average, people reported that they themselves were more malleable than others, and self-theories were a better predictor of students' own personal academic motivation and responses (De Castella & Byrne, 2015). Likewise, Aneeta Rattan and colleagues demonstrated that people may not apply the same theory of mutability to all people or groups. People who believe that the capacity for improvement is universal are more likely to support policies that promote equal opportunity, while those who believe that only some people have the capacity to become highly intelligent are less inclined to support such measures (Rattan & Georgeac, this volume; Rattan, Savani, Naidu, & Dweck, 2012).

advantage by shifting people to an incremental (high-effort, improvement-focused) mindset at times of failure, which is when this lay theory is particularly important. In an academic domain, it is possible that successful individuals subtly shift toward a belief in their stable, enduring skills to build confidence when doing well, but readily switch to a belief in mutability and improvement when they encounter setbacks.

One thing, we have noted across the relatively few existing studies is that there is a considerably stronger tendency to shift to an incremental theory when failure is encountered than to shift to an entity theory when focused on a personal success (Leith et al., 2014; Steimer & Mata, 2016). Theoretically, this is consistent with the view that people are more likely to respond in a motivated manner when faced with threat. It may be that when outcomes are favorable to the self, there is no motive to recruit information that selectively supports a particular conclusion: an individual can enjoy a success whether they believe that the capacity is fixed or changeable. Nonetheless, we suspect that under certain conditions, people may be especially motivated to shift toward an entity perspective following success. This intuition is consistent with the belief of parents that praising ability is beneficial for self-esteem (Mueller & Dweck, 1998), and reflected in people's tendency to attribute positive (but not negative) personal outcomes to dispositional qualities (Campbell & Sedikides, 1999). There is something gratifying about the idea that one's successes come from within, and reflect some enduring set of qualities that can be counted on to continue panning out in the future. We suspect that these self-esteem benefits are at least part of the reason people so readily shift their implicit theories when provided with praise about their abilities (Mueller & Dweck, 1998)—it feels good. However, an overreliance on this entity perspective quickly becomes counterproductive if it shifts people away from mastery toward performance goals, and prompts helpless responses to failure (Hong et al., 1999). We offer some speculations regarding who may be most likely to actively shift their lay theories toward an entity perspective after encountering success.

Our first speculation is based on the notion that some successes feel more fragile than others. Sometimes, we can see the clear path from our time and effort to a desirable outcome. In these cases, we may feel confident that we can control similar successful outcomes in the future, and an incremental theory might be just as gratifying as an entity theory, and there would be little motivation to shift the dominant theory. On the other hand, we sometimes encounter successes that we are not so confident we can reproduce. This may occur when hard work and outcome are not so obviously causally related; in other words when success is experienced as noncontingent on performance—a circumstance that leads people to self-handicap (Jones & Berglas, 1978). It may occur in those settings that tend to produce the “imposter syndrome,” (possibly especially prevalent amongst high-achieving women, Clance & Imes, 1978) where people have difficulty taking credit for the accomplishments or kudos they have garnered and worry that outcomes were based on luck or some other circumstance not controlled by themselves. It may also occur when people's internal lack of confidence (low self-esteem, for instance) leads them to view positive outcomes as inconsistent with expectations. When people encounter

these successes but worry that they may be fleeting and out of their control, we reason that one response might be to shift to an entity theory in an effort to psychologically “stake a claim” to the abilities that presumably underlie their successes. We have no direct evidence that these conditions are especially likely to prompt motivated adoption of entity theories, though there is some correlational evidence that the experience of the imposter syndrome is linked to entity beliefs about capacity in women (Kumar & Jagacinski, 2006). We also reported some very preliminary observations about responses to success in Leith et al. (2014, Footnote 6). We wondered if the natural ups and downs of academic life might contribute to people’s shifting lay theories over the course of a semester. We measured students’ implicit theories (intelligence and general person) at the beginning of Fall semester, then followed up in the Winter semester (4–6 months later). We asked people to report at Time 2 on the outcomes they regarded as disappointments and successes. We found that students who reported a greater proportion of disappointments over the previous semester showed a slight tendency to become more incremental in their views, whereas students who reported a greater proportion of successes over the semester showed a significant shift toward an entity theory. We interpret these results with considerable caution due to a small sample size ($N = 41$) and the exploratory nature of the work, but suggest that the shift toward an entity theory among students (perhaps tenuous) experiences of greater success may reflect a desire to feel that their recent accomplishments will bode well for their future outcomes. Given that a strong entity theory appears to have considerable downsides for individuals’ motivation and achievement (Burnette et al., 2010), we suggest that it is worthwhile to develop a better understanding of the factors that lead people to actively adopt these fixed beliefs.

Our second speculation pertains to how context may contribute to shifts toward either an incremental or entity mindset more generally, but where a particular set of risk factors for adopting too strong an entity theory may emerge. We suspect that many people go through life encountering a pretty robust mix of successes and failures: even those who work hard and demonstrate notable success are likely to take on bigger and bigger challenges, which sometimes will lead to setbacks and failures. However, some people are likely to find themselves in contexts where one type of outcome is especially likely to occur a majority of the time. For instance, gifted students may not only be at risk, as Dweck (2012) suggests, of being frequently praised for their intelligence, they are also likely to find themselves in situations where academic successes far outweigh failures, giving them few opportunities to incorporate failure and the capacity for improvement into their self-views and beliefs. One of this chapter’s authors (Wilson) has noted this entity inclination not infrequently in incoming graduate students: often these students have had a preponderance of past experiences as the best and brightest scholars in their cohort. Graduate school offers even the most talented students a host of opportunities for setbacks and failures, which can initially be quite a shake-up for students’ self-views. Wilson has taken to delivering informal incremental “interventions” at times of setback, hoping to “strike while the iron is hot” and trigger lay theory change when students may be especially motivated to shift.

Another context where entity implicit theories may be prompted is in children's sports. Children who demonstrate high performance at particular sports are often plucked out of recreational leagues and recruited for more elite teams. In some cases, the "best of the best" are combined into teams that typically outcompete most others in their category. These kids may go seasons at a time without every encountering the experience of losing a game. Although the hard work of athletes and their coaches is a fundamental part of success even on these teams, the players on these "superteams" may come to think of their ability as inborn, and struggle once they find themselves moving up to a level of competition where they once again face losing. We speculate that some players who have been encouraged through experience to cultivate an entity view of athletic ability might be especially likely to worry that they "don't have what it takes" when they progress to the next level of athletic challenge. Indeed, in 2006, Carol Dweck was asked to develop a training intervention with the Blackburn Rovers, a soccer team in the United Kingdom's premium league (Krakovsky, 2007), when their coach expressed concern about how a "star is born not made" mentality was keeping very good players from reaching their full potential. These talented players were stuck in an entity mindset, believing that inborn ability would carry the day, and hence neglecting their rigorous training schedule. Dweck designed an intervention starting with the youngest and most impressionable players, fostering an incremental mindset to instill a belief in the value of effort and training.

How Person-Perception Goals May Shape Implicit Theories of Stability and Change

Although people arguably spend a good deal of time thinking about themselves, they also spend a significant proportion of their time observing, interacting with and perceiving others. Sometimes we simply want to get an accurate impression of a new person in order to predict our likely future interactions with them. However, in other cases, we have a vested interest in how information is processed about important others. We may interpret a close friend's foibles—say forgetting to return the clothing she borrowed—as endearing or accidental. Conversely, we might be inclined to view equally ambiguous actions of a disliked ex-spouse—say forgetting to update a scheduling conflict—as malicious and intentional. We thought that when individuals are invested in seeing particular others' temporally extended actions in either a favorable or unfavorable light, they may show an inclination to gravitate toward the implicit theory that supports the conclusion they prefer to draw. In other words, by activating an incremental theory, the best friend can always become more responsible the next time she borrows clothes; by adopting an entity theory we can assure ourselves that the ex-spouse will never change.

We suspect that these motives would play out for any kind of close relationship. In an initial test of these hypotheses, we focused on public figures for

whom participants would have a vested interest in either excusing or disparaging. We chose Canadian and American political figures who, over a period of years, represented their respective political party (Leith et al., 2014, Studies 4–6). For example, in the run-up to a Federal election taking place in Canada in 2011, we asked people about the Liberal candidate Michael Ignatieff and the Conservative candidate Stephen Harper. We presumed that people who affiliated with one of these political parties would be motivated to view their candidate in a favorable light and to view the opposing candidate less magnanimously. During this election, both candidates had taken some criticism for statements they had uttered years before, which now cast them in an unflattering light. For example, Ignatieff, often critiqued for insufficient patriotism, was quoted as having called the Canadian flag a “passing imitation of a beer label.” Stephen Harper, critiqued for a lack of empathy, was quoted as having said “In terms of the unemployed, of which we have over a million-and-a half, I don’t feel particularly bad for many of these people.” We collected a set of unflattering past utterances by both candidates an average of 10 years prior, and randomly assigned liberal and conservative participants to read them. As we expected, people’s beliefs about the changeable nature of these candidates was highly contingent on participants’ political stripes. Conservatives were certain Harper was, at core, a changeable person but Ignatieff’s qualities were hopelessly fixed. Liberals demonstrated precisely the same convictions—but about the opposite candidate. Further, believing the candidate was changeable mediated people’s belief that these decade-old foibles were simply not relevant to their current judgment of the politician; an entity view, on the other hand, supported the belief that those past missteps were highly pertinent to judgments of political character today. Of course, the idiosyncratic wrongs of the two candidates were not easily comparable; in follow-up studies with greater experimental control we described candidates’ political past as either poor or commendable (for instance, describing Barack Obama’s time in Senate as earning him an overall A or C grade from a bipartisan review committee). After reading about a poor Senate record, Republicans viewed Obama as more fixed and unchangeable than did Democrats who viewed him as highly changeable. Republicans, on the other hand, saw Obama as far more changeable after success than they had seen him to be after failure.

One shortcoming of these two previously described studies is that we measured a very specific lay theory—how changeable one particular politician was believed to be. This arguably diverges from the notion that lay theories guide more general information processing. In a follow-up study, we tested the logic more fully: we presented favorable or unfavorable information about politicians (this time Justin Trudeau and Stephen Harper in Canada), and then asked participants about their *general* person lay theories, such as “*People can do things differently sometimes, but the important parts of who they are can’t really be changed.*” When given the opportunity to endorse a sweeping lay theory as it applies to people in general, political affiliation still guided which theory they were inclined to endorse. Participants who read about their favored candidates’ foibles believed that *people in general* can change more than those who read about their candidate’s accomplishments; the reverse was true for the opposing candidate.

Judging the relevance of the past for the present is an ambiguous task we are faced with in many spheres of life. For instance, Americans were recently faced with the question of whether Donald Trump's 2005 recording in which he bragged about kissing and groping women without their consent is reflective of who he is today. In an apology video, Trump said: "Anyone who knows me, knows these words don't reflect who I am." Rudy Giuliani, too, invoked an incremental view in an interview about this incident: "That was then and this is now. And he's gone through 14 months of running for president. And, as you know, running for president does something to you. It changes a lot of the way you look at things, it changes a lot of the way in which you behave." Reminded of his own past infidelities, Giuliani further endorsed a general theory of malleability, saying: "We believe that people in this country can change." (ABC News, 2016, Oct 9). In the same interview, Donna Brazile (Chair of the Democratic National Committee) countered: "This is not a changed man. This is who Donald Trump truly is." This kind of temporally extended judgment can be daunting: does that past action signify a lasting clue to a person's character? Has the person learned and grown from a past mistake, becoming even wiser and more trustworthy as a result? It makes sense that we would draw on our beliefs about the fundamental nature of people's change and stability to answer these questions. However, less obviously, we suggest that when we draw on these lay theories, *which* lay theory we choose to endorse at that moment may be plucked out of our array of beliefs because it will best help us to reach a particular conclusion. We have the experience of reasoning about the situation by drawing on our knowledge of typical human mutability, and may not be particularly aware—or concerned—that these mutability beliefs shift from one context to the next.

We have also begun to think about other contexts in which the dominant lay theory activated during a judgment can have meaningful consequences for other important outcomes. For example, judgments of the appropriate way to approach crime and punishment dependents deeply on one's beliefs about the possibility of rehabilitation. Not surprisingly, if an entity theorist espouses the view that "once a thief, always a thief," their judgment—and recommended punishment—of an offender may be considerably more harsh than an incremental theorist who believes any past transgressor can "turn over a new leaf." Indeed, Gervy, Chiu, Hong, and Dweck (1999) found that entity theorists were more likely to value principles of punishment over rehabilitation, while incremental theorists put more weight in rehabilitation over punishment. When lay theories are conceived as chronic individual differences, we might understand people's beliefs about the fundamental mutability of criminal offenders' moral status as a basic philosophical perspective which informs their views of crime and punishment. We wondered if lay theories may be subject to greater change than previously assumed even in these contexts. Again, we began with the premise that the context would have to prompt a motivation to shift one's lay theories to reach a desired conclusion. For example, Todd may have a punitive stance on criminals in general, believing that people's basic moral character never changes. However, in the event that his son is arrested, he might quickly begin to recruit knowledge of how changeable people's moral

foundations may be—that sometimes, people just need a second chance to learn from their mistakes.

We reasoned that a variety of motives could be relevant to judgments of crime: our judgments of loved ones might be clouded by generosity, and our assessments of outgroup members clouded by prejudice or mistrust. In Study 7 of Leith et al. (2014), we examined how people might shift their implicit theories of how changeable people are at their core after reading about a serious criminal offender (someone who had been convicted and served time for child sexual assault). We recognized that recidivism beliefs would be particularly high across the board for such a crime, so to increase variability in judgments, we described in detail evidence of the offender's rehabilitation. Next, we considered what kinds of factors would produce a motivated judgment in such a case. We reasoned that parents would be especially concerned about protecting their children, prompting additional vigilance when faced with this type of offender. We also reasoned that the physically "closer to home" the offender was seen to be, the more motivated people would be to protect their family. How might this vigilance be reflected in respondents' endorsement of implicit theories? We reasoned that the most threatened group (parents who considered a nearby offender) would be motivated to stay wary and keep their guard up by presuming that people do not change their basic qualities. This would allow them to remain mistrustful of the indications of rehabilitation and would support their opposition to the offender's placement. To test these ideas, we recruited parents and nonparents and asked them to consider the (hypothetical) case in which this offender, out on parole, requests relocation to a city 200 miles away from them, or relocation into the participants' own community. We then asked respondents whether people, in general, can change their core characteristics. The group of participants that we expected to be most threatened, thus motivated to shift their implicit theories, were parents who imagined the offender in their own community. As we expected, those respondents ignored evidence of rehabilitation and reported the strongest conviction that people simply cannot change their basic attributes. Of course, we recognize that there are evidence-based differences in the likely recidivism rates of different types of crimes, and we do not argue that this information is irrelevant. What we point out, however, is that information *other than* evidence can shape people's beliefs about the likelihood of mutability and therefore rehabilitation—a motivated process with highly consequential outcomes.

We have begun in recent research to investigate other contexts that might motivate shifts in lay theories of change and stability (Williams & Wilson, 2016). In keeping with our focus on crime and punishment, we wondered whether people—particularly those who are high in prejudice—might shift in their lay theories when judging criminals of different races. In an initial test of this hypothesis, we asked participants to read a news article about an offender who had committed a crime some time in the past. Race (Caucasian/African American) was subtly varied by using name (e.g., DeShawn vs. Bradley) as a cue. We found that people high (but not low) in prejudice toward African Americans were more likely to shift to an entity theory when they read about an African American offender, relative to when

the offender was depicted as White. These entity views again had consequences: they mediated harsher punishment recommendations for the crime.

Does Motivation Guide the Adoption of Other Lay Theories of Mutability?

Although Carol Dweck's research on personal beliefs about the mutability of attributes has received widespread attention, these are not the only lay theories about the dynamic or fixed nature of human attributes. We point to two other sets of lay theories that, amongst other features, contain assumptions about immutability or change. One closely linked literature focuses on genetic essentialism and the belief that various characteristics, behaviors or conditions are genetically determined (Dar-Nimrod & Heine, 2011; Haslam, this volume). The other set of beliefs that share features of an incremental theory refers to people's understanding of how society functions: beliefs in social mobility, meritocracy, and the "American Dream" (e.g., Kraus & Tan, 2015).

Genetic Theories

Like entity theorists, people who believe that an attribute is genetically determined tend to view outcomes as more immutable. Although genetic essentialism carries with it other beliefs as well (e.g., about etiology), the mutability beliefs overlap very closely with Dweck's approach to implicit theories. For instance, believing intelligence or body size is genetically determined is akin to having an entity theory of intelligence or weight. However, the literature on genetic essentialism has been more explicitly grounded in public and scientific discussion and debate around topics such as intergroup differences and social inequality, whereas the implicit theories literature has been characterized as occurring in more of a "social vacuum" (Jayaratne et al., 2006).

People vary, for example, in their belief that racial or sex characteristics are genetic (and hence, group characteristics are immutable), fuelling a debate about whether unequal group outcomes are due to inherent factors or due to social context and opportunity (Jayaratne et al., 2006). Jayaratne and colleagues reported that genetic accounts of racial differences tends to be linked to higher levels of racial prejudice, though from the correlational design it cannot be established whether genetic theories foster racism or whether racism motivates endorsement of genetic theories. The authors suggest that the process is likely bidirectional, strongly pointing to a genetic lay theory as a "legitimizing myth" that has historically justified prejudice and discriminatory practices. At the same time, they note that genetic lay theories may become prevalent for non-motivated reasons—for

instance, the rise of genomics and behavioral genetic research—which can influence or reinforce people’s entity beliefs about groups.

It is also the case that politically conservative (and upper class) individuals are more likely to endorse the genetic roots of racial and class differences (Kraus & Keltner, 2013; Suhay & Jayaratne, 2013), a process that the authors suggest may also be motivated (see also Hegarty & Golden, 2008). Suhay and Jayaratne suggest that with various causal attributions available in media and public discourse, individuals can “pick and choose” the explanations that best allow them to support their ideological and social position. These divergent explanations of group differences are also reflected in the media: conservative newspapers contain more biological explanations for sex differences than more liberal newspapers (Brescoll & LaFrance, 2004). The authors argue that this difference in emphasis of one causal theory or another allows conservatives to recruit the ideological underpinning that justify the status quo, while allowing liberals to identify sociocultural explanations to support a desire to change the existing system. The belief that group differences are inborn also appears to increase when people are threatened, supporting a motivated account. For example, people are often motivated to justify the system in which they live, even when it produces injustices. Activating system-justification motives increases people’s endorsement of an essentialist and immutable view of sex differences (Brescoll, Uhlmann, & Newman, 2013; Morton, Postmes, Haslam, & Hornsey, 2009). Similarly, Morton, Hornsey and Postmes (2009) found that prejudiced people appeal to an essentialist view of race when the outcome would exclude an outgroup, but de-essentialize race when the outcome would exclude their ingroup.

Suhay and Jayaratne (2013) also demonstrate the striking flexibility of people’s endorsement of genetic lay theories. Although conservatives invoke genetic accounts for perceived race or class differences (e.g., intelligence, aggression, etc.) more than liberals, liberals, and conservatives do not differ in their genetic explanations for these same characteristics as possessed by *individuals*. Further, the endorsement of genetics flips when providing an account for a different stigmatized group: gay men and lesbians. Here, liberals are more apt to argue that people are born with a particular sexual orientation (because emphasizing lack of choice and inability to change delegitimizes moral approbation), and conservatives are more likely to point to context, upbringing, and “lifestyle choice.” Much of the documented link between genetic attributions and prejudice has been correlational, hence, there has been debate about whether a belief that homosexuality is innate drives acceptance (Brewer, 2008), or whether, instead, increasing societal acceptance of gay rights has motivated supporters to adopt a genetic view (Lewis, 2009). Recent research offers some support for a motivated reasoning account: people are more likely to be influenced by information that supports the causal attribution (genetic or environmental) that aligns with their political viewpoints (Morin-Chassé, Suhay, & Jayaratne, 2014; Suhay & Garretson, 2015). In other words, their ideology appears to influence the lay theory they adopt more than their lay theory affects their ideology. Once again, we do not claim to assess the validity of any given theory of nature versus nurture—some are almost certainly more

correct than others. We instead highlight how the availability of both lay theories in public discourse allows people to choose the viewpoint that best justifies their values or prejudices.

Implicit Theories of Social Mobility

Western—and perhaps especially American—society remains highly committed to notions of social mobility and meritocracy, even as conditions of increasing inequality have made this belief less and less a reality (Hacker & Pierson, 2010; Kraus & Tan, 2015; Piketty, 2014). Obama (2012) characterized the American Dream—which he believed was under siege—as “the basic American promise that if you worked hard, you could do well enough to raise a family, own a home, send your kids to college, and put a little away for retirement.” The American dream has at its core the very incremental idea that by working hard and applying sufficient effort, anyone can get ahead. When we consider this societal-level myth rather than the individual incremental beliefs that Dweck (2012) so strongly recommends, some of the pitfalls of an overly incremental theory become evident. There is evidence that belief in meritocracy and social mobility can increase people’s tolerance of societal inequality (Larsen, 2016; Manza & Brooks, 2016; Shariff, Wiwad, & Aknin, 2016), and that strong meritocracy beliefs lead people to overlook the fact that, at least in American society, inequality of opportunity limits the degree to which meritocracy can fairly allocate outcomes (Hacker, 2006). The American Dream has been implicated in why people may vote for policies that work against their own interest—for example, why working-class people would support tax cuts going disproportionately to the wealthy. Belief in the equalizing power of hard work, lower socioeconomic status (SES) individuals can justify a system that has prevented them from realizing their ever-extolled American Dream. Given, the puzzlingly ways in which these beliefs work against people’s own interests, we suspected that there may be strong motivations to cling to an incremental belief in social mobility despite evidence to the contrary.

Kraus and Tan (2015) directly address this paradox in their work on social mobility. People tend to overestimate, in general, the likelihood of someone rising up in social class. This exaggerated belief in the American Dream myth may simply be due to the cultural prevalence of these ideas. However, Kraus and Tan also suggest that people may be motivated to cling to these beliefs. Specifically, when people were asked to estimate the social mobility of someone *similar to themselves*, belief in mobility increased significantly. Notably, belief in the malleable nature of social status may be self-serving for both the rich and poor; believing in the flexible nature of social classes allows rich people to justify their status as earned through hard work (Kraus, Davidai, & Nussbaum, 2015; Kraus & Keltner, 2013; Kraus & Tan, 2015). In turn, belief in mobility offers hope and alleviates threat for those less well off (Davidai & Gilovich, 2015). Justifying their system by believing in the power of hard work and the American dream allows for the reduction of dissonance

and the acceptance of blatant social inequality (Jost, Pelham, Sullivan, & Sheldon, 2003). Indeed, the authors theorized (and found) that low-income participants demonstrated a stronger belief in the legitimacy of social inequality and were more likely to support the statement that “large differences in pay are necessary to foster motivation and effort” (Jost et al., 2003). This finding is parallel to Leith et al.’s (2014) findings for individual failure experiences; low-income people who feel the sting of failure to rise in status may gravitate to the incremental view that it is still, nonetheless, possible. Paradoxically, the motivation to resolve this dissonance can cause those who suffer the most from these social inequalities to justify the status quo that keeps them in a low status position (Jost, Banaji, & Nosek, 2004).

The ardent—and perhaps motivated—belief in the link between hard work and success may also underlie a tendency to blame the poor for their own outcomes. If everyone can get ahead, why haven’t they? As Du Monteil (2015) argues, “That’s the whole idea of the American Dream: only those who work hard for it, are hungry for it, and don’t give up in face of adversity are actually able to live it.” The corollary assumption, of course, is that those who remain poor must just have not tried hard enough. This perception of the undeserving poor overlaps considerably with another lay belief: the conviction that the world is just and fair. Just World Theory posits that people are motivated to believe that the world is a just place, where people get what they deserve (Hafer & Bègue, 2005; Lerner, 1980). People with a stronger Belief in a Just World (BJW) are particularly likely to endorse both social mobility and meritocracy (Day & Fiske, 2016), a constellation of beliefs that would all allow them to conclude that the poor are to blame for their own fate. Researchers have also experimentally demonstrated that people make judgements reflective of a stronger BJW (e.g., rating the poor as less intelligent) after exposure to evidence of injustice and inequality in society (Kay, Jost, & Young, 2005), presumably due to their system-justification motivation. In another study, Iatridis and Fousiani (2009) asked participants to read about a student with either high or low socioeconomic status (SES) who encountered either academic success or failure. Participants explained the high-SES student’s success in terms of ability and the low-SES student’s success as luck, whereas when they read about failure they thought the high-SES student had not exerted enough effort and the low-SES student did not have enough ability. Further, participants endorsed a higher BJW when the high-SES student succeeded and the low-SES student failed. Intriguingly, a meta-analysis by Malahy, Rubinlicht, and Kaiser (2009) examined whether actual levels of inequality observed in America between 1973 and 2006 were related to average levels of BJW identified from studies conducted during that time span which included the measure. They found that as income inequality in the USA has risen, so has Americans’ endorsement that the world is just and that people get what they deserve. Malahy et al. interpret this pattern of BJW as potentially reflective a motivated, system-justifying response to the injustice inherent in an increasingly unequal society, and caution that the belief may inhibit empathy for the plight of the disadvantaged and decrease support for programs intended to redistribute or foster equality of opportunity.

Conclusions and Future Directions

We have considered a number of contexts in which people's assumptions about human and societal change may be more fluid than often supposed, and how the assumptions people adopt in a given context can underlie—almost invisibly—their consequential judgments about social policy, about other individuals, and about themselves. One particularly interesting—but insidious—aspect of people's beliefs about change is that they may often provide the foundation for people's subsequent judgments, yet the beliefs themselves go unexamined and undiscussed. Even when core assumptions about human change are expressed, it is difficult to definitively determine who is factually correct when it comes to the nature of human mutability. As a result, people may often be puzzled by those who offer strikingly divergent judgments of the same action, because they base their judgments on different purported “truths” about human nature.

The literature systematically examining motivated shifts in people's lay theories is still limited, and a goal of this chapter is to encourage further inquiry. First, we suggest that the relative impact of chronicity and fluidity is not well understood. People do appear to have chronic lay beliefs that guide their everyday information processing, sometimes leading them astray. However, we have identified a number of contexts where motivated reasoning likely influences the adoption of one theory or another. We do not yet know how much of human behavior is best represented by chronic differences in beliefs or by flexibly shifting assumptions. The mounting evidence of the fluidity of lay theories might even call into question the assumption of chronic individual differences—if people have knowledge of both theories and can activate one or the other depending on the context and their goals, then some evidence for chronic lay theories may actually be due to chronic contextual factors and motives. We are certainly not ready to disregard the notion that people tend to have a dominant theory that guides them in the absence of factors that could prompt them to change those views. Indeed, we suspect that there are also individual differences in the degree to which people fluidly shift from one theory to another. Some people may flexibly adopt the theory that best supports their preferred conclusions; others may find themselves stuck in a mindset that works against their interests in some contexts.

We call for further research examining these questions, as well as the downstream consequences of motivated shifts in lay theories. We argue that lay theories have meaningful real-world consequences; it may be that the consequences produced when people engage in motivated shifting may contribute to their longer term dominance of one theory or another. For instance, if, after failure, people adopt an incremental theory, they may persist at the task more effectively and actually improve. As a result, they will have accumulated evidence for malleability, which may in turn reinforce a chronic incremental mindset.

We also do not know whether a lay theory shift in one domain may influence subsequent judgments in another domain—for instance, if I shift to endorse an entity theory to impugn a political candidate whose past actions I still revile, might I

subsequently be stuck with that entity theory if asked to make judgments another candidate—or about a downstream outcome such as crime and punishment? Some of these questions also reflect a lack of precision in our knowledge of the mechanism (what cognitive process leads people’s lay theories to shift) and in people’s level of awareness of these shifts (are people conscious of shifting theories when they do it?). We suggest that answers to these questions will not only contribute to the new area of inquiry regarding the fluidity of lay beliefs, but also build a more nuanced understanding of the ways that individuals actively construct their more chronic beliefs over time.

References

- ABC News. (2016, October 9). *Transcript of this week with George Stephanopoulos*. Retrieved from <http://abcnews.go.com/Politics/week-transcript-rudy-giuliani-donna-brazile/story?id=42670926>
- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113–125. doi:10.1006/jesp.2001.1491
- Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680–740). New York, NY: McGraw-Hill.
- Blackwell, L., Trzesniewski, K., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246–263.
- Bradfield, J. P., Taal, H. R., Timpson, N. J., et al. (2012). A genome-wide association meta-analysis identifies new childhood obesity loci. *Nature Genetics, 44*, 526–531.
- Brescoll, V. L., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on gender differences. *Psychological Science, 15*, 515–521.
- Brescoll, V. L., Uhlmann, E. L., & Newman, G. E. (2013). The effects of system-justifying motivations on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology, 105*, 891–908.
- Brewer, P. R. (2008). *Value war: Public opinion and the politics of gay rights*. New York: Rowman & Littlefield.
- Burnette, J. L., & Finkel, E. J. (2012). Buffering against weight gain following dieting setbacks: An implicit theory intervention. *Journal of Experimental Social Psychology, 48*, 721–725. doi:10.1016/j.jesp.2011.12.020
- Burnette, J. L., Pollack, J. M., & Hoyt, C. L. (2010). Individual differences in implicit theories of leadership ability and self-efficacy: Predicting responses to stereotype threat. *Journal of Leadership Studies, 3*, 46–56. doi:10.1002/jls.20138
- Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology, 3*, 23–43. doi:10.1037/1089-2680.3.1.23
- Chiu, C., Hong, Y., & Dweck, C. S. (1997a). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology, 73*, 19–30. doi:10.1037/0022-3514.73.1.19
- Chiu, C. Y., Dweck, C. S., Tong, J. Y., & Fu, J. H. (1997b). Implicit theories and conceptions of morality. *Journal of Personality and Social Psychology, 73*, 923–940.
- Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice, 15*, 241–247. doi:10.1037/h0086006

- Davidai, S., & Gilovich, T. (2015). Building a more mobile America—One income quintile at a time. *Perspectives on Psychological Science, 10*, 60–71.
- Dar-Nimrod, I., & Heine, S. (2011). Genetic essentialism: on the deceptive determinism of DNA. *Psychological Bulletin, 137*, 800–818.
- Day, M. V. & Fiske, S. T. (2016). Movin' on up? How perceptions of social mobility affect our willingness to defend the system. *Social Psychological and Personality Science, 8*, 267–274. doi:10.1177/1948550616678454.
- De Castella, K., & Byrne, D. (2015). My intelligence may be more malleable than yours: the revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education, 30*, 245–267. doi:10.1007/s10212-015-0244-y
- Du Monteil, E. (2015). *20 reasons to believe in the American Dream*. XPat Nation. <http://xpatnation.com/reasons-to-believe-in-the-american-dream/>
- Dweck, C. S. (2012). Implicit theories. In P. Van Lange, A. Wruglanski, & E. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 43–62). London: SAGE Publications Limited.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995a). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry, 6*, 267–285. doi:10.1207/s15327965pli0604_1
- Dweck, C. S., Chiu, C., & Hong, Y. (1995b). Implicit theories: Elaboration and extension of the model. *Psychological Inquiry, 6*, 322–333. doi:10.1207/s15327965pli0604_12
- Dweck, C. S., Hong, Y., & Chiu, C. (1993). Implicit theories individual differences in the likelihood and meaning of dispositional inference. *Personality and Social Psychology Bulletin, 19*, 644–656. doi:10.1177/0146167293195015
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273. doi:10.1037/0033-295X.95.2.256
- Gervey, B. M., Chiu, C., Hong, Y., & Dweck, C. S. (1999). Differential use of person information in decisions about guilt versus innocence: The role of implicit theories. *Personality and Social Psychology Bulletin, 25*, 17–27.
- Gladwell, M. (2013). *David and Goliath: Underdogs, misfits, and the art of battling giants* (1st ed.). New York, NY: Little, Brown and Company.
- Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development, 84*, 1526–1541. doi:10.1111/cdev.12064
- Hacker, J. (2006). *The great risk shift: The new economic insecurity and the decline of the American dream*. New York, NY: Oxford University Press.
- Hacker, J., & Pierson, P. (2010). *Winner-take-all politics: How Washington made the rich richer and turned its back on the middle class*. New York, NY: Simon and Schuster.
- Hafer, C. L., & Bègue, L. (2005). Experimental research on just-world theory: Problems, developments, and future challenges. *Psychological Bulletin, 131*, 128–167.
- Harris, J. A., Vernon, P. A., Olson, J. M., & Jang, K. L. (1999). Self-rated personality and intelligence: A multivariate genetic analysis. *European Journal of Personality, 13*, 121–128.
- Haslam, N. (2017). The origins of lay theories: The case of essentialist beliefs. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Hegarty, P., & Golden, A. M. (2008). Attributional beliefs about the controllability of stigmatized traits: Antecedents or justifications of prejudice? *Journal of Applied Social Psychology, 38*, 1023–1044.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.
- Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M. S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*, 588–599. doi:10.1037/0022-3514.77.3.588
- Hong, Y., Levy, S. R., & Chiu, C. (2001). The contribution of the lay theories approach to the study of groups. *Personality and Social Psychology Review, 5*, 98–106. doi:10.1207/s15327957PSPR0502_1

- Iatridis, T., & Fousiani, K. (2009). Effects of status and outcome on attributions and just-world beliefs: How the social distribution of success and failure may be rationalized. *Journal of Experimental Social Psychology, 45*, 415–420. doi:10.1016/j.jesp.2008.12.002
- Jayaratne, T. E., et al. (2006). White Americans' genetic lay theories of race differences and sexual orientation: Their relationship with prejudice toward blacks, and gay men and lesbians. *Group Processes Intergroup Relations, 9*, 77–94. doi:10.1177/1368430206059863
- Jones, E. E., & Berglas, S. (1978). Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement. *Personality and Social Psychology Bulletin, 4*, 200–206. doi:10.1177/014616727800400205
- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25*, 881–919.
- Jost, J. T., Pelham, B. W., Sheldon, O., & Sullivan, B. N. (2003). Social inequality and the reduction of ideological dissonance on behalf of the system: Evidence of enhanced system justification among the disadvantaged. *European Journal of Social Psychology, 33*, 13–36. doi:10.1002/ejsp.127
- Kammrath, L., & Dweck, C. S. (2006). Voicing conflict: Preferred conflict strategies among incremental and entity theorists. *Personality and Social Psychology Bulletin, 32*, 1497–1508.
- Kay, A. C., Jost, J. T., & Young, S. (2005). Victim derogation and victim enhancement as alternate routes to system justification. *Psychological Science, 16*, 240–246.
- Kelly, L. E. (1955). Consistency of the adult personality. *American Psychologist, 10*, 659–681.
- Kelly, A. M., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex, 15*, 1089–1102.
- Knowles, E. D., & Ditto, P. H. (2012). Preference, principle, and political casuistry. In J. Hanson (Ed.), *Ideology, psychology, and law* (pp. 341–379). New York, NY: Oxford University Press.
- Krakovsky, M. (2007). The effort effect. *Stanford Magazine*. https://alumni.stanford.edu/get/page/magazine/article/?article_id=32124
- Kraus, M. W., Davidai, S., & Nussbaum, A. D. (2015, May 3). American dream? Or mirage? *The New York Times*.
- Kraus, M. W., & Keltner, D. (2013). Social class rank, essentialism, and punitive judgment. *Journal of Personality and Social Psychology, 105*, 247–261.
- Kraus, M. W., & Tan, J. J. X. (2015). Americans overestimate social class mobility. *Journal of Experimental Social Psychology, 58*, 101–111.
- Kumar, S., & Jagacinski, C. M. (2006). Imposters have goals too: The imposter phenomenon and its relationship to achievement goal theory. *Personality and Individual Differences, 40*, 147–157. doi:10.1016/j.paid.2005.05.014
- Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53*, 636–647.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480–498. doi:10.1037/0033-2909.108.3.480
- Larsen, C. A. (2016). How three narratives of modernity justify economic inequality. *Acta Sociologica, 59*, 93–111.
- Levy, S. R., Chiu, C., & Hong, Y. (2006). Lay theories and intergroup relations. *Group Processes & Intergroup Relations, 9*, 5–24. doi:10.1177/1368430206059855
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology, 74*, 1421–1436. doi:10.1037/0022-3514.74.6.1421
- Leith, S. A., Ward, C. L. P., Giacomini, M., Landau, E. S., Ehrlinger, J., & Wilson, A. E. (2014). Changing theories of change: Strategies shifting in implicit theories endorsement. *Journal of Personality and Social Psychology, 107*, 597–620.
- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Lewis, G. B. (2009). Does believing homosexuality is innate increase support for gay rights? *The Policy Studies Journal, 37*, 669–693.

- Lou, N. M., & Noels, K. A. (2016). Changing language mindsets: Implications for goal orientations and responses to failure in and outside the second language classroom. *Contemporary Educational Psychology, 46*, 22–23. doi:[10.1016/j.cedpsych.2016.03.004](https://doi.org/10.1016/j.cedpsych.2016.03.004)
- Malahy, L. W., Rubenlicht, M. A., & Kaiser, C. R. (2009). Justifying inequality: A cross-temporal investigation of U.S. income disparities and just-world beliefs from 1973 to 2006. *Social Justice Research, 22*, 369–383.
- Manza, J., & Brooks, C. (2016). *Prisoners of the American Dream? Americans' Attitudes Towards Inequality and Taxes in the New Gilded Age*. Manuscript under review.
- Molden, D. C., & Dweck, C. S. (2006). Finding “meaning” in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist, 61*, 192–203. doi:[10.1037/0003-066X.61.3.192](https://doi.org/10.1037/0003-066X.61.3.192)
- Molden, D. C., Plaks, J. E., & Dweck, C. S. (2006). “Meaningful” social inferences: Effects of implicit theories on inferential processes. *Journal of Experimental Social Psychology, 42*, 738–752. doi:[10.1016/j.jesp.2005.11.005](https://doi.org/10.1016/j.jesp.2005.11.005)
- Morin-Chassé, A., Suhay, E., & Jayaratne, T. (2014). Discord over DNA: Politically contingent responses to scientific research on genes and race. APSA 2014 Annual Meeting Paper; American University School of Public Affairs Research Paper No. 2014-0002.
- Morton, T. A., Hornsey, J. M., & Postmes, T. (2009a). Shifting ground: The variable use of essentialism in contexts of inclusion and exclusion. *British Journal of Social Psychology, 48*, 35–59. doi:[10.1348/014466607X270287](https://doi.org/10.1348/014466607X270287)
- Morton, T. A., Postmes, T., Haslam, S. A., & Hornsey, M. J. (2009b). Theorizing gender in the face of social change: Is there anything essential about essentialism? *Journal of Personality and Social Psychology, 96*, 653–664.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children’s motivation and performance. *Journal of Personality and Social Psychology, 75*, 33–52. doi:[10.1037/0022-3514.75.1.33](https://doi.org/10.1037/0022-3514.75.1.33)
- Nussbaum, A. D., & Dweck, C. S. (2008). Defensiveness versus remediation: Self-theories and modes of self-esteem maintenance. *Personality and Social Psychology Bulletin, 34*, 599–612.
- Obama, B. (2012). *State of the Union Address*.
- Peez, J., & Wilson, A. E. (2008). The temporally extended self: The relation of past and future selves to current identity, motivation, and goal pursuit. *Social and Personality Psychology Compass, 2*, 2090–2106.
- Peez, J., & Wilson, A. E. (2014). Marking time: Selective use of temporal landmarks as barriers between current and future selves. *Personality and Social Psychology Bulletin, 40*, 44–56. doi:[10.1177/0146167213501559](https://doi.org/10.1177/0146167213501559)
- Piketty, T. (2014). *Capital in the twenty-first century*. Cambridge, MA: Belknap Press.
- Plaks, J. E., Levy, S. R., & Dweck, C. S. (2009). Lay theories of personality: Cornerstones of meaning in social cognition. *Social and Personality Psychology Compass, 3*, 1069–1081. doi:[10.1111/j.1751-9004.2009.00222.x](https://doi.org/10.1111/j.1751-9004.2009.00222.x)
- Plaks, J. E., & Stecher, K. (2007). Unexpected improvement, decline, and stasis: A prediction confidence perspective on achievement success and failure. *Journal of Personality and Social Psychology, 93*, 667–684. doi:[10.1037/0022-3514.93.4.667](https://doi.org/10.1037/0022-3514.93.4.667)
- Poon, C. S. K., & Koehler, D. J. (2006). Lay personality knowledge and dispositionist thinking: A knowledge-activation framework. *Journal of Experimental Social Psychology, 42*, 177–191. doi:[10.1016/j.jesp.2005.04.001](https://doi.org/10.1016/j.jesp.2005.04.001)
- Poon, C. S. K., & Koehler, D. J. (2008). Person theories: Their temporal stability and relation to intertrait inferences. *Personality and Social Psychology Bulletin, 34*, 965–977. doi:[10.1177/0146167208316690](https://doi.org/10.1177/0146167208316690)
- Pyszczynski, T., & Greenberg, J. (1987). Self-regulatory perseverance and the depressive self-focusing style: A self-awareness theory of reactive depression. *Psychological Bulletin, 102*, 122–138. doi:[10.1037/0033-2909.102.1.122](https://doi.org/10.1037/0033-2909.102.1.122)
- Rattan, A., & Dweck, C. S. (2010). Who confronts prejudice? The role of implicit theories in the motivation to confront prejudice. *Psychological Science, 21*, 952–959.

- Rattan, A., & Georgeac, O. (2017). Mindsets about malleability and intergroup relations. In C. M. Zedelius, B. C. N. Mülle, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Rattan, A., Savani, K., Naidu, N. V. R., & Dweck, C. S. (2012). Can everyone become highly intelligent? Cultural differences in and societal consequences of beliefs about the universal potential for intelligence. *Journal of Personality and Social Psychology*, *103*, 787–803. doi:[10.1037/a0029263](https://doi.org/10.1037/a0029263)
- Reis, H. T. (2008). Reinvigorating the concept of situation in social psychology. *Personality and Social Psychology Review*, *12*, 311–329. doi:[10.1177/1088868308321721](https://doi.org/10.1177/1088868308321721)
- Ross, M., & Wilson, A. E. (2002). It feels like yesterday: Self-esteem, valence of personal past experiences, and judgments of subjective distance. *Journal of Personality and Social Psychology*, *82*, 792–803.
- Sasaki, J. Y., LeClair, J., West, A. L., & Kim, H. S. (2016). The gene-culture interaction framework and implications for health. In J. Y. Chiao, S. Li, R. Seligman, & R. Turner (Eds.), *The oxford handbook of cultural neuroscience* (p. 279). New York, NY: Oxford University Press.
- Schumaker, J. A., & Slep, A. M. S. (2004). Attitudes and dating aggression: A cognitive dissonance approach. *Prevention Science*, *5*, 231–243. doi:[10.1023/B:PREV.0000045357.19100.77](https://doi.org/10.1023/B:PREV.0000045357.19100.77)
- Sedikides, C. (1993). Assessment, enhancement, and verification determinants of the self-evaluation process. *Journal of Personality and Social Psychology*, *65*, 317–338. doi:[10.1037/0022-3514.65.2.317](https://doi.org/10.1037/0022-3514.65.2.317)
- Shariff, A. F., Wiwad, D., & Akin, L. B. (2016). Income mobility breeds tolerance for income inequality: Cross-national and experimental evidence. *Perspectives on Psychological Science*, *11*, 373–380.
- Shepperd, J., Malone, W., & Sweeny, K. (2008). Exploring causes of the self-serving bias. *Social and Personality Psychology Compass*, *2*, 895–908. doi:[10.1111/j.1751-9004.2008.00078.x](https://doi.org/10.1111/j.1751-9004.2008.00078.x)
- Steimer, A., & Mata, A. (2016). Motivated implicit theories of personality: My weaknesses will go away, but my strengths are here to stay. *Personality and Social Psychology*, *42*, 415–429. doi:[10.1177/0146167216629437](https://doi.org/10.1177/0146167216629437)
- Skitka, L. J., Bauman, C. W., & Mullen, E. (2008). Morality and justice: An expanded theoretical perspective and empirical review. In K. A. Hegtvedt & J. Clay-Warner (Eds.), *Advances in group processes: Justice*.
- Suhay, E., & Garretson, J. (2015, March). Science, sexuality, and civil rights: Does research on the causes of homosexuality have a political impact? 8th Annual NYU-CESS Experimental Political Science Conference.
- Suhay, E., & Jayaratne, T. E. (2013). Does biology justify ideology? The politics of genetic attribution. *Public Opinion Quarterly*, *77*, 497–521.
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*, 755–769. doi:[10.1111/j.1540-5907.2006.00214.x](https://doi.org/10.1111/j.1540-5907.2006.00214.x)
- Tesser, A. (2001). On the plasticity of self-defense. *Current Directions in Psychological Science*, *10*, 66–69. doi:[10.1111/1467-8721.00117](https://doi.org/10.1111/1467-8721.00117)
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1130. doi:[10.1126/science.185.4157.1124](https://doi.org/10.1126/science.185.4157.1124)
- Ward, C., & Wilson, A. E. (2015). Implicit theories of change and stability moderate effects of subjective distance on the remembered self. *Personality and Social Psychology Bulletin*, *41*, 1167–1179.
- Williams, S. L., & Wilson, A. E. (2016, January). *Once a thief, always a thief? How time, implicit theories, and race affect moral judgments*. Poster presented at the annual meeting of the Society for Personality and Social Psychology, San Diego, California, USA.
- Wood, J. V., Giordana-Beech, M., Taylor, K. L., Michela, J. L., & Gaus, V. (1994). Strategies of social comparison among people with low self-esteem: Self-protection and self-enhancement. *Journal of Personality and Social Psychology*, *67*, 713–731. doi:[10.1037/0022-3514.67.4.713](https://doi.org/10.1037/0022-3514.67.4.713)

- Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 302–314. doi:[10.1080/00461520.2012.722805](https://doi.org/10.1080/00461520.2012.722805)
- Yeager, D. S., Trzesniewski, K. H., & Dweck, C. S. (2013). An implicit theories of personality intervention reduces adolescent aggression in response to victimization and exclusion. *Child Development, 84*, 970–988. doi:[10.1111/cdev.12003](https://doi.org/10.1111/cdev.12003)
- Ziegler, A., & Stoeger, H. (2010). A learning oriented subjective action space as an indicator of giftedness. *Psychological Science Quarterly, 50*, 222–236.

Part II
Explorations in Lay Theories About
Human Psychological Attributes or
Phenomena

Lay Theories of Self-control

Veronika Job and Gregory M. Walton

People sometimes fail to behave as they intend. Children who participated in the famous marshmallow experiments wanted to wait and not to eat a tempting marshmallow placed on the plate before them. They knew that they would get a second marshmallow if they managed to wait until the experimenter came back. Still, many children ate the one marshmallow, losing the chance to get a second one (Mischel, Shoda, & Rodriguez, 1989). Often people give into temptations and immediate impulses at the cost of their long-term goals. Aggression and violence, over-eating, impulsive spending, and sexual behavior, as well as drug addiction, are examples of failures in self-control (Baumeister & Heatherton, 1996; Gottfredson & Hirschi, 1990; Vohs & Faber, 2007). They can have disastrous consequences for an individual and high cost for society as a whole. Therefore, considerable psychological research has aimed to understand self-control failures and investigate ways to improve self-control. The research reviewed in this chapter focused on people's beliefs about the availability of self-control capacity as an important predictor of their self-control.

During the last few decades, social-psychological research on self-control has been dominated by a model that uses a simple metaphor to explain self-control failures: the strength model of self-control (Baumeister, Bratslavsky, Muraven, & Tice 1998; Muraven, Tice, & Baumeister, 1998). This model suggests that self-control relies on a limited resource. Engaging in self-control draws down this resource, leaving the individual with reduced capacity to exert further self-control. In support for this model, studies have found that after people have exerted

V. Job (✉)

Department of Psychology, University of Zurich,
Binzmuehlestrasse 14/11, 8050 Zurich, Switzerland
e-mail: v.job@psychologie.uzh.ch

G.M. Walton

Department of Psychology, Stanford University,
450 Serra Mall, Stanford, CA 94305, USA

self-control on one task subsequent self-control performance is impaired, even on a very different kind of task. This is termed the *ego depletion effect*. The strength model of self-control stimulated an impressive array of empirical research. Ego depletion effects were documented for outcomes as diverse as intellectual performance (Schmeichel, Vohs, & Baumeister, 2003), information processing (Fischer, Greitemeyer, & Frey, 2008), impression management (Vohs, Baumeister, & Ciarocco, 2005), and resisting violent responses to a partner's provocation (Finkel, DeWall, Slotter, Oaten, & Foshee, 2009) (for meta-analytic summaries and controversies, see Carter, Kofler, Forster, McCullough, 2015; Hagger, Wood, Stiff, & Chatzisarantis, 2010). Indeed, ego depletion has been described as a universal phenomenon based in physiology (Gailliot & Baumeister, 2007).

However, increasing evidence questions this model. New research finds that the ego depletion effect is not inevitable but is affected by motivation (e.g., Inzlicht & Schmeichel, 2012; Inzlicht, Schmeichel, & Macrae, 2014; Molden et al., 2012) and other moderators (for overviews see Loschelder & Friese, 2016; Masicampo, Martin, & Anderson, 2014). For instance, monetary incentives, autonomy, and positive mood can prevent ego depletion effects (Moller, Deci, & Ryan, 2006; Muraven & Slessareva, 2003; Tice, Baumeister, Shmueli, & Muraven, 2007). Further, people's expectancies about their ability to exert self-control following the exertion of self-control can moderate ego depletion. People were told that performing an effortful task (controlling their emotions) could either improve or harm performance on a subsequent task (Martijn, Tenbült, Merckelbach, Dreezens, & de Vries, 2002). Participants' subsequent self-control performance confirmed their expectations: Those who expected self-control depletion performed worse, while those who expected a self-control boost performed better.

We ask a more general question. Perhaps, it is not just that motivation helps people overcome depletion. Perhaps lay theories about self-control in general give rise to ego depletion in the first place (see also Mukhopadhyay & Johar, 2005). People may draw from society's general theories about the nature of self-control capacity. These general theories are expressed, for example, in cultural products like movies or advertisements that echo and promulgate a specific belief ("You're not you when you're hungry"). Holding a global theory that difficult tasks deplete one, making it difficult to sustain self-control efforts, may make people feel depleted, exhausted, and in need of rest and replenishment when they face high demands. We expected that people who do not endorse this limited theory on self-control—who instead believe that self-control efforts can even become self-energizing—may not experience depletion.

In this chapter we provide an overview of research on lay theories of self-control. Laboratory research shows that these theories, both measured as an individual difference and manipulated to examine causality, predict performance as people take on a series of self-control tasks. The theory that self-control does not rely on a limited resource helps people sustain self-control performance. Further, theories of self-control matter in everyday life settings. They predict self-regulatory success as well as well-being, specifically when people face high demands.

Measuring Lay Theories of Willpower

A first step to investigating people’s lay theories on self-control was to develop a measure. Previous research showed that lay theories (e.g., of the malleability of personal attributes) are typically domain specific (Dweck, 1999). Since we expected the same for lay theories of self-control, we decided to begin by focusing on strenuous mental activities. We developed a scale containing six items that reflected the belief that self-control, for which we also use the colloquial term willpower, relies on a limited and easily depleted resource (*limited-resource theory*; e.g., “After a strenuous mental activity your energy is depleted and you must rest to get it refueled again”). Reverse-coded items (e.g., “After a strenuous mental activity, you feel energized for further challenging activities”) referred to the opposite belief: They reject the idea that willpower is highly constrained and, instead, suggest that exerting willpower can be energizing. We called this belief a *nonlimited* theory about willpower (Table 1). People with a nonlimited theory do not necessarily believe that self-control capacity is infinite or that they can exert self-control endlessly, never needing to rest or sleep. It is not an “unlimited” belief. However, they reject the view that willpower is readily depleted by acts of self-control.

Depending on the research question and the purpose of a study, we have developed additional scales assessing lay theories in other domains, like resistance to temptation, emotion control, and physical exertion (Bernecker & Job, 2015a, in press). The domain-specific scales represent distinct factors that best predict specific behavior when matched to the behavior in question. For example, in a study conducted with Type 2 diabetes patients, only willpower theories in the domain of

Table 1 Items to measure implicit theories about willpower for strenuous mental activity (Job et al., 2010)

Strenuous mental activity	
1. Strenuous mental activity exhausts your resources, which you need to refuel afterward (e.g., through taking breaks, doing nothing, watching television, eating snacks)	R
2. After a strenuous mental activity, your energy is depleted and you must rest to get it refueled again	R
3. When you have been working on a strenuous mental task, you feel energized and you are able to immediately start with another demanding activity	
4. Your mental stamina fuels itself. Even after strenuous mental exertion, you can continue doing more of it	
5. When you have completed a strenuous mental activity, you cannot start another activity immediately with the same concentration because you have to recover your mental energy again	R
6. After a strenuous mental activity, you feel energized for further challenging activities	

R reversed items

resisting temptations predicted junk food eating and only willpower theories in the domain of strenuous physical activity predicted physical activity (Bernecker & Job, 2015a).

The willpower theories scales we have developed so far are certainly not exhaustive and future research may well develop measures in other self-control domains.

Do People Experience Ego Depletion Only if They Believe that Willpower Relies on a Limited Resource?

Our first set of studies tested whether ego depletion is “all in your head,” that is, whether it occurs only if people *believe* that willpower relies on a limited resource (Job, Dweck, & Walton, 2010).

In a first study, we tested the hypothesis that peoples’ habitual lay theories about willpower, as assessed with the “strenuous mental activity” scale, would moderate the ego depletion effect. After filling out the scale, participants completed a difficult self-control task that has been used in previous research to induce a state of depletion, or an easier version of the task that does not require self-control. We then assessed a classic laboratory measure of self-control: Stroop performance. Color words (red, green, yellow, and blue) appeared on a computer screen in one of the four colors. Participants were instructed to indicate the color of the font, which either matched or did not match the meaning of the word. This task requires self-control because on incongruent trials (e.g., the word “red” displayed in green) people have to suppress the meaning of the word. Previous research has found ego depletion effects on performance on these trials (Inzlicht, McKay, & Aronson, 2006; Webb & Sheeran, 2003). Participants who held a limited-resource theory showed the same pattern. After completing the difficult rather than easy initial task, they made more mistakes on incongruent trials on the Stroop task. But participants who held a nonlimited theory performed equally well whether they had completed a difficult “depleting” or nondepleting task first. These results have been replicated several times; by us (Job, Walton, Bernecker, & Dweck, 2013) and independently (Chow, Hui, & Lau, 2015; Salmon, Adriaanse, De Vet, Fennis, & De Ridder, 2014).

Research also extends these findings to other self-control domains, including resistance to temptation (Bernecker & Job, in press). In one study, participants who had to resist a tempting food (freshly baked buns) (see Baumeister et al., 1998) later showed poor Stroop performance to the extent that they endorsed the theory that people have a limited capacity to resist temptations. Illustrating the domain-specificity of willpower theories, people’s theories about whether strenuous mental activities depend on a limited resource did not predict performance in this study (the two scales were only modestly correlated). Another study examined theories about emotion control (Bernecker & Job, in press, Study 2). We

manipulated whether people had to suppress their emotions during a funny video (see Brown & McConnell, 2011). Lay theories about emotion control (but not strenuous mental activity) predicted how long people persisted on a subsequent frustrating task. People who thought that their capacity to control emotions depends on a limited resource persisted less long when they had had to suppress their emotions than when they had not. But people who thought the capacity to control emotions does not depend on a limited resource showed no decrement in persistence when they had had to suppress their emotions.

Taken together, these studies suggest that people who hold nonlimited willpower theories do not experience ego depletion in a wide range of self-control domains. Moreover, they illustrate that theories of willpower are domain specific. If people think that strenuous mental activity drains a limited resource, they become depleted after such a task and show impaired self-control even in other domains (e.g., Stroop). However, they do not necessarily feel depleted after resisting a temptation, since they do not necessarily believe that resistance to temptations depends on a limited and depletable resource. To matter, the implicit theory assessed in a research study has to match the initial, “depleting” task.

In the studies described so far, theories about willpower were only measured. They thus do not demonstrate their causal effect. Perhaps people who actually have greater self-control in a domain see this self-control as less limited, and this is why they can maintain their self-control performance as self-control demands accumulate. To address possibilities like this and test the causal effects of willpower theories, we manipulated theories about willpower using two biased questionnaires containing easy-to-agree-with items that endorsed either a limited or a nonlimited-resource theory (e.g., “Working on a strenuous mental task can make you feel tired such that you need a break before accomplishing a new task” vs. “Sometimes, working on a strenuous mental task can make you feel energized for further challenging activities,” see Table 2). This procedure evokes thoughts consistent with a specific theory. Manipulation checks consistently show that in each condition participants agreed with the suggested theory. As predicted, in the context of strenuous mental activity, participants led to endorse the limited-resource theory showed ego depletion (Job et al., 2010). They performed worse after they had completed a “depleting” task as compared to the nondepleting control task. In contrast, participants led to endorse the nonlimited theory showed no drop in self-control performance following the initial “depleting” task as compared to an undemanding task. Thus, manipulated theories about willpower showed parallel effects on self-control performance as measured theories.

An artificial aspect of most laboratory research on self-control including the above-mentioned studies is that they assess performance on a single self-control task rather than sustained effort and success over time. A further study tested whether willpower theories might predict people’s sustained efforts over the course of a challenging learning task (Miller et al., 2012). Participants completed the biased questionnaire that manipulated their theory about willpower. They then engaged for 20 min with a demanding task widely used to improve working

Table 2 Biased questionnaires to manipulate implicit theories about willpower (Job et al., 2010)

Induction of limited-resource theory
1. When you think over a matter with great concentration, it can be sometimes tiring
2. Working on a strenuous mental task can make you feel tired much so that you need a break before accomplishing a new task
3. When you have to do many demanding activities for a while, you eventually get exhausted and less productive
4. Sometimes, when you completely focus your attention on a demanding mental activity, you feel tired and you need a break sooner or later since your resources have to be refilled
5. After you have been working on a strenuous mental task for several hours, you can get fatigued so that you need to rest before taking on the next challenging activity
6. Strenuous mental activity sometimes exhausts your resources, which you need to refuel afterward (e.g., through breaks, doing nothing, watching television, eating...)
7. After a strenuous mental activity, your energy can be depleted and you sometimes must rest to get it refuelled again
8. Sometimes, when you have completed a very exhausting mental activity, you have to recover your mental energy again before starting with the same concentration on a new difficult task
Resisting temptations
1. Sometimes, it can be very inspiring to think over a matter with great concentration
2. When situations accumulate that challenge you with temptations, it gets more and more difficult to resist the temptations
3. It can be energizing to be completely focused on a demanding mental activity, so that you are able to remain concentrated for a while
4. Sometimes, it is energizing to be fully absorbed with a demanding mental task
5. It can be energizing to be completely focused on a demanding mental activity, so that you can remain concentrated for a long time
6. Sometimes, your mental stamina fuels itself. After a strenuous mental exertion you can continue doing more of it
7. It is possible to be in such a productive work mode that you do not need much recreation between different mentally strenuous tasks
8. Working on a strenuous mental task can activate your mental resources and you become even better at accomplishing subsequent demanding tasks

memory. Although all participants learned effectively at the beginning of the task, improving their performance over the first half of the task, participants led to think of willpower as a limited resource stalled over the second half of the task, whereas those in the nonlimited theory condition sustained their learning and improvements in performance over the full task.

These laboratory experiments indicate that reduced performance after previous self-control exertion (i.e., “ego depletion”) results not from a true lack of resources but from people’s beliefs about their self-control capacity. They raise important questions. First, many goals in the real-world demand sustained self-control. Can implicit theories about willpower affect the extent to which people accomplish their

goals and flourish in their daily lives? And second, what mechanisms underlie the effects of willpower theories on sustained self-control?

Effects of Lay Theories About Willpower in Everyday Life

The strength model of self-control suggests that understanding the limits of self-control capacity should help people use their limited resources wisely and therefore predict better self-regulation and well-being. People with a nonlimited theory, by contrast, may overuse their resources and suffer from severe depletion when high demands accumulate (Vohs, Baumeister, & Schmeichel, 2012). However, in contrast to this view, we hypothesized that a limited-resource theory leads people to let up on self-control efforts long before they have reached any actual limit. If so, they may reduce their effort on everyday tasks and fail to accomplish their goals, especially when they face high demands. But people with a nonlimited theory might better sustain their efforts, improving their everyday self-regulation.

Everyday Self-regulation

In a first longitudinal study (Job et al., 2010, Study 4), we tracked college students across three time points over an academic term, the last of which was during final exams. The results showed that only students who endorsed a limited-resource theory (assessed with the strenuous mental activity and the resisting temptations scale) self-regulated less effectively at the stressful final time point, for instance reporting procrastinating more and eating more junk food controlling for baseline self-regulation.

This study simply assumed that self-regulatory demands were high for all students as final exams approached. A second study examined the level of self-regulatory demands each student reported on a week-by-week basis so as to distinguish students who faced high demands from those who faced lower demands (Job, Walton, Bernecker, & Dweck, 2015). As predicted, although students with limited and nonlimited willpower theories faced similar levels of self-regulatory demands, only those with a limited theory showed increasing self-regulation failures (e.g., procrastination, junk food eating, bad time management) as demands increased. In addition, this study assessed students' end-of-term grade point average (GPA). Among students who took a heavy course load, students with the non-limited theory earned higher grades than students with the limited theory.

A recent study extended these findings to people with diabetes, who face particularly high and significant self-control challenges (Bernecker & Job, 2015a). To control their blood sugar levels, people with diabetes have to adhere to a complex regimen involving regular blood sugar testing, medication, a low-glycemic diet, and

exercise throughout their entire life (Boule, Haddad, Kenny, Wells, & Sigal, 2001; Brand-Miller, Petocz, Hayne, & Colagiuri, 2003). Would lay theories about willpower predict how well patients adhere to their therapy? In a correlational study, type 2 diabetes patients completed three theories about willpower scales (resisting temptations, physical activity, strenuous mental activity) and measures assessing therapy adherence (i.e., blood glucose monitoring, diet, exercise) and psychological adjustment (i.e., emotional distress, well-being, life quality). As predicted, participants with a limited theory reported fewer self-care activities (full scale), a less healthy diet (resisting temptations scale), and less physical activity (physical activity scale) than people with a nonlimited theory. They also reported more emotional distress from the disease and experienced less subjective well-being and reported a worse life quality. The belief that willpower is nonlimited seems to be more adaptive for coping with the demands that arise from managing diabetes than the belief that willpower is limited.

Personal Goal-Striving and Well-Being

Another important context of self-regulation involves the personal goals people set for themselves, and their success in accomplishing them. For instance, a person may have the goal to be admitted to a specific college or to lose ten pounds of weight. Personal goals are conscious representations of anticipated end-states. They represent what people strive for and want to achieve in life (Emmons, 1986; Klinger, 1977; Little, 1983). Several theorists propose that having goals and striving for and achieving them is crucial for the development and maintenance of well-being, because goals provide meaning, structure, and direction to a person's life as well as, when completed, a sense of accomplishment (Brunstein, 1993; Diener, Suh, Lucas, & Smith, 1999; Emmons, 1986; Maier & Brunstein, 2001).

In early research on personal goal-striving, Mukhopadhyay and Johar (2005) showed that measured and manipulated beliefs about self-control as depending on a limited or a nonlimited resource and as either malleable versus fixed affected the number of New Years' resolutions people set. People who thought that self-control is not limited set more New Years' resolutions. Moreover, a second study showed that manipulating lay theories about willpower affected people's success in the keeping of their resolutions. Participants led to view willpower as dependent on a limited resource were less likely to succeed 4 months later, especially if they had set difficult goals. It seems that they were more likely to give up in the face of difficulties or setbacks.

To examine more directly whether lay theories of self-control predict goal attainment most when demands accumulate, one of the above-mentioned longitudinal studies (Job et al., 2010, Study 4) assessed self-regulation with respect to a personal achievement goal. At the first assessment period, students listed a personal goal that involved challenge and achievement. People were asked at each subsequent time point over the term how well they had regulated themselves in pursuing

this goal (e.g., “I was often not in the mood to do something for this goal”). As predicted, during the demanding final exam week, students who had a limited theory about willpower reported worse goal-related self-regulation than students with a nonlimited theory. Another study found that with a limited-resource theory even demands experienced on a day-to-day basis can undermine self-regulatory efforts toward personal goals (Bernecker & Job, 2015b).

If people with a nonlimited theory make more progress toward their personal goals, do they experience greater well-being? In another set of studies, we first found a strong relationship between theories about willpower and life satisfaction as well as affective well-being (Bernecker, Herrmann, Brandstätter, & Job, 2017). The more people endorsed a nonlimited theory about willpower, the greater was their subjective well-being. Next, a longitudinal study tested whether willpower theories predicted change in subjective well-being over students’ first year in college. As expected, a limited theory about willpower predicted a decline in subjective well-being from a period with low demands (i.e., the beginning of the first year) to a period with high demands (i.e., final exams at the end of the first year). Another longitudinal study replicated this finding using a daily diary method and, moreover, showed that the gains in well-being for people with a nonlimited theory of willpower were mediated by more effective goal-striving and more progress toward personal goals over the course of the term (Bernecker et al., 2015, Study 3). These findings show that a nonlimited theory of willpower does not just help people accomplish tasks in the face of demands. It does not just make people better workers. It helps people accomplish goals—people’s own priorities for their lives—and this improves their well-being.

Mechanisms

How does a limited theory about willpower undermine people’s efforts at self-control, especially as demands accumulate? So far, several potential mechanisms have been explored: perceived exhaustion, sensitivity to cues about the availability of resources, activation of the goal to rest, and self-efficacy.

Perceived Exhaustion

The first evidence that perceived exhaustion may play an important factor came from a series of experiments conducted by Clarkson and colleagues, which found that a manipulation of the exhaustion people perceived in a previous task affected subsequent self-control performance, whereas actual self-control exertion did not (Clarkson, Hirt, Jia, & Alexander, 2010). Does a limited theory about willpower make people experience self-control exertion as more exhausting, and is this what reduces subsequent performance? To test this question, we assessed perceived

exhaustion (“How exhausting was the task?”) in a study manipulating implicit theories about willpower (Job et al., 2010, Study 3). Even though the manipulation affected subsequent self-control performance, it had no effect on the degree to which people perceived the “depleting” task as exhausting (see also Job, Bernecker, Miketta, & Friese, 2015). Instead, willpower theories affected how people responded to the experience of exhaustion. People in the limited willpower theory condition responded to feelings of exhaustion with decrements in their subsequent self-control performance. The more exhausting they found the first task, the worse they performed on the second task. This was not the case for people in the non-limited theory condition. Although they found the “depleting” task just as exhausting, for them, feelings of exhaustion were not a reason to let up on their self-control efforts. A limited-resource theory seems to attune people to experiences of exhaustion, and to take this as a sign to let up.

Sensitivity to Cues About the Availability of Resources

If people with a limited-resource theory are sensitive to perceived exhaustion, are they sensitive to cues about the availability of mental resources more generally? This resource sensitivity hypothesis is supported by another set of studies, which link theories about willpower to the finding that ingested glucose, too, buffers the ego depletion effect (Job et al., 2013).

Previous research showed that ingesting glucose can improve self-control performance and buffer ego depletion (DeWall, Baumeister, Gailliot, & Maner, 2008; Gailliot et al., 2007). Research suggests, however, that ingested glucose does not simply restore depleted energy resources and directly fuel performance (Kurzban, 2010). In contrast, peripheral sensory receptors in the mouth and digestive system, which are sensitive to glucose, can activate reward regions in the brain and increase motivation (Chambers, Bridge, & Jones, 2009; Kringsbach, 2004; Kurzban, 2010). Merely rinsing the mouth with glucose, as compared to a sugar substitute, improves physical performance and mitigates ego depletion (Hagger & Chatzisarantis, 2013; Molden et al., 2012; Sanders, Shirk, Burgin, & Martin, 2012). Thus, people who have ingested glucose may perform better because these peripheral cues signal the availability of energy, motivating them to sustain effort on difficult tasks. If a limited-resource theory sensitizes people to cues about the availability of resources, then theories about willpower may moderate the effect of glucose on subsequent self-control performance.

Three experiments found evidence for this hypothesis (Job et al., 2013). Replicating past research, people who reported holding a limited-resource theory, or who were induced to hold this theory, showed improved self-control performance following an initial demanding task when they had consumed glucose (lemonade with sugar) rather than a substitute (lemonade with a sugar substitute). Yet, people with a nonlimited theory showed no such benefit. They performed well regardless of whether they consumed the sugar or the nonsugar drink. This was the case even

though participants could not reliably distinguish the sugar from the nonsugar drink in their self-reports. The results provide further evidence that self-control does not rely on a limited physiological resource that is depleted by even brief acts of self-control and is restored by glucose consumption (Gailliot & Baumeister, 2007; Gailliot et al., 2007; see Kurzban, 2010; Molden et al., 2012). Instead, a limited-resource theory attunes people to cues about the availability of resources, including cues below conscious awareness. They further document how top-down beliefs interact with bottom-up physiological information to influence people's self-regulatory success.

Activation of a Rest Goal

If a limited-resource theory attunes people to cues to their internal states (perceived exhaustion) and the availability of self-control resources (glucose), does it also activate the goal to rest following self-control exertion? This hypothesis is consistent with the process model of self-control (Inzlicht & Schmeichel, 2012; Inzlicht et al., 2014), which proposes that after people exert self-control, they are no longer motivated to exert themselves and this is why people perform worse on subsequent self-control tasks.

Because it is well documented that people possess limited introspective abilities that often lead to invalid self-reports about inner motivational states (e.g., Silvia & Gendolla, 2001; Wilson & Dunn, 2004) a series of studies assessed motivational shifts after “depletion” using indirect implicit and behavioral indicators, including reaction times (RTs), object evaluations, and actual resting (Job, Bernecker, Miketta, & Friese, 2015; Job, Walton, Bernecker, & Dweck, 2015).

In one study, we tested whether people with a limited-resource theory would value means to reach the goal to rest more strongly once they were “depleted”. We assumed that after engaging in a self-control task they would evaluate objects that are helpful for resting (bed, sofa, hammock, cup of tea, bathtub, TV screen) more positively (see Ferguson & Bargh, 2004; Fishbach, Shah, & Kruglanski, 2004). After reporting their theories about willpower and completing the depletion manipulation, participants were asked how much they liked both objects relevant to rest and objects relevant to physical or mental exertion (barbell, racing bicycle, punching bag, treadmill, sneakers, Sudoku puzzles). As predicted, in the high-“depletion” condition, the more people endorsed the limited-resource theory, the more highly they evaluated rest-conducive objects, and the more they devalued objects conducive to physical and mental exertion. Willpower theories were not related to evaluations in the low depletion condition. After self-control exertion, a limited-resource theory both inclines people to value rest and recovery and disinclines activities that involve effort and exertion. A second study found the same result after manipulating theories about willpower, confirming their causal effect (Job, Bernecker, Miketta, & Friese, 2015, Study 3).

If people with a limited-resource theory want to rest, do they rest more if given the chance? In additional studies, a limited-resource theory—both measured and manipulated—led people to rest longer following a “depleting” experience before continuing

with another task. In one study, only participants randomly assigned to a limited theory condition and to a high depletion condition took an excessively long time to complete an ostensible “product-tasting” task in which they could lounge in comfortable chairs following the depletion task (Job, Bernecker, Miketta, & Friese, 2015, Study 5).

These data show that the belief that willpower depends on a limited resource causes a motivational shift toward rest following the exertion of self-control. The findings are consistent with the process model of self-control, which denies the existence of a specialized self-control resource and explains ego depletion effects through shifts in motivation and attention (Inzlicht & Schmeichel, 2012; Inzlicht et al., 2014). That model postulates that after having expended effort in a strenuous task, people are less motivated to expend further effort. Our research adds an important specification: Only people who think, or are led to think, that self-control relies on a limited resource show the motivational shift toward rest.

Changes in Self-efficacy

A recent line of research suggests changes in self-efficacy as a further mechanism underpinning the effects of lay theories of willpower on self-control. Self-efficacy is the “judgment of how well one can execute courses of action required to deal with prospective situations” (Bandura, 1982, p. 122). People tend to prefer to engage effort in tasks that they perceive themselves to be good at and to withdraw from tasks that seem difficult to them. Chow et al. (2015) proposed that when people exert self-control their self-efficacy for upcoming tasks is temporarily reduced, which impairs further performance. Moreover, they suggest, this reduction in self-efficacy occurs only in people with a limited-resource theory. People with a nonlimited theory about willpower should not react to self-control exertion with reduced self-efficacy because for them exerting self-control does not imply a lack of available resources (Chow et al., 2015).

Three experiments supported this theorizing. First, they showed that people depleted by an initial challenging self-control task reported reduced self-efficacy to exert further self-control. A second study confirmed that this reduction in self-efficacy mediated the effect of depletion on subsequent self-control performance. Finally, a third experiment confirmed that only people with a limited theory about willpower showed this drop in self-efficacy following self-control exertion. Moreover, the drop in self-efficacy mediated the moderating effect of a limited willpower theory on subsequent self-control performance (Chow et al., 2015).

Relations Among Mechanisms

So far different research lines explored three different mechanisms explaining why a limited theory about willpower leads to reduced self-control when demands

accumulate. It is likely that these processes interact. For instance, the inference that cues (e.g., feelings of exhaustion, lack of sleep, time since last snack) signal a lack of needed resources sets off two motivational shifts: 1) a feeling of reduced self-efficacy (“I can’t do more”) and the activation of a rest goal (“I want to rest”). Thus there may be a reciprocal relationship between these processes dragging down people’s willingness and perceived ability to exert self-control. It is a task for future research to examine these concurring processes and to integrate them into a comprehensive model.

Boundaries and Possible Negative Consequences

The findings we have presented thus far suggest that a nonlimited theory about willpower is more beneficial than a limited theory about willpower, both in laboratory self-control tasks and in everyday self-regulation, goal-striving, and well-being. Could a nonlimited theory be counterproductive in some circumstances?

Overuse of Resources

Vohs et al. (2012) hypothesized that a nonlimited theory may lead people to “overuse” resources, temporarily compensating for depleted resources, and thus improve self-control performance in the face of mild or moderate self-control demands but not in the face of high demands. In a laboratory experiment, they compared a “no depletion” condition (no initial self-control tasks), a “mild depletion” condition (two initial self-control tasks), and a “severe depletion” condition (four initial self-control tasks). Theories about willpower were manipulated with the biased questionnaire. First, the study replicated our previous findings: In the “mild depletion” condition participants led to think of willpower as a nonlimited resource sustained a high level of performance. But in the “severe depletion” condition, there was no positive effect of a nonlimited theory. Moreover, on one of two measures of self-control performance, the effect reversed. Participants in the severe depletion condition performed worse when they had been led to think of willpower as nonlimited. Vohs and colleagues concluded that a nonlimited theory can be counterproductive. Thinking that willpower is nonlimited, they suggest “might undermine the normal tendency to conserve resources so that people find themselves severely depleted after multiple tasks” (Muraven, Shmueli, & Burkley, 2006, p. 186).

As a laboratory session wears on, however, many factors beyond people’s self-control capacity may affect their willingness to exert further effort on tasks of little personal relevance. People in the severe depletion condition may simply have been unwilling to exert further effort on such tasks, regardless of their willpower theory. Indeed, a nonlimited theory about willpower would not be functional if it led people to engage on a high level with every task that came along regardless of its value or purpose. Future research may distinguish the capacity to exert self-control

from the value or meaning of a task to the self, for instance by comparing tasks of personal relevance to those without. However, from our perspective the critical test of the functionality of willpower theories comes from field studies examining people's efforts to accomplish their own goals in their daily lives. As discussed earlier, examining students' self-regulatory success in a demanding academic environment was predicted by a nonlimited theory, especially when they faced the greatest demands (Job, Walton, Bernecker, & Dweck, 2015). Further, among people with type 2 diabetes who face high and significant self-regulatory demands, the nonlimited theory predicted greater therapy adherence (Bernecker & Job, 2015a).

Of course, it is possible that alternative processes may arise when people face extreme physical or psychological circumstances (e.g., torture). As we have emphasized, a nonlimited theory is not an unlimited theory. The belief that willpower is not limited does not imply that people think they can continue to control themselves and exert effort indefinitely without needing to rest, sleep, or eat. What research on implicit theories shows is that, as compared to the belief that willpower relies on a limited resource, the belief that willpower does not rely on a limited resource simply helps people stay engaged for longer during the normal range of challenges they face in their daily lives. Put the other way: the limited-resource theory undermines people's self-control success by leading people to reduce effort and conserve their "resources" long before they reach any true limits.

Interpersonal Consequences

Most research on theories of willpower has examined people's efforts to accomplish their own goals. But if a person believes that willpower does not rely on a limited resource, do they expect more not only from themselves but also from others? If other people fail to meet these expectations, are nonlimited theorists less understanding and harsher in their judgments? Although not examining limited-resource beliefs, one line of research found that peoples' beliefs about willpower as either malleable or a fixed trait (measured and manipulated) predicted harsher judgments of people with salient self-regulatory failures (e.g., to quit smoking, to lose weight; Freeman, Shmueli, & Muraven, 2013). However, high expectations can also be helpful in promoting people's performance—when these expectations are communicated in positive, growth-oriented ways (e.g., Cohen, Steele, & Ross, 1999; Lepper, Woolverton, Mumme, & Gurtner, 1993). Future research can examine how individuals can communicate a nonlimited theory about willpower in ways that support and improve other people's self-control.

Consequences for Parenting

Mukhopadhyay and Yeung (2010) examined how lay theories about self-control affect parenting. They theorized that parents who think of willpower as not reliant on a limited resource would not sufficiently prioritize the development of self-control skills in their children. They reasoned that “the belief that reserves of self-control are already large may lessen the value of further developing these reserves” (p. 242). Accordingly, they expected that parents with a limited theory about willpower, who in addition believe that the limited capacity can be enlarged (limited, but malleable theorists), would engage more in behaviors that help develop children’s self-control as compared to parents with a nonlimited theory. Indeed, a series of studies showed that parents who believed that willpower relies on a limited resource but is malleable were more likely to restrict unhealthy snacking and fast-food consumption in their children as compared to nonlimited-malleable theorists. They were further more likely to choose educational television programs for their children. A manipulation of theories about self-control (nonlimited vs. limited/fixed vs. limited/malleable) further confirmed their causal effect. Adults who led to believe that self-control does not rely on a limited resource were more likely to choose gifts for a child that provided instant pleasure. But adults led to believe that self-control is limited but malleable chose gifts that were more educational. They were further convinced that their choice would have a positive effect on the child’s development. People with a nonlimited theory did not emphasize the development of children’s self-control in their choice. Apparently, they did not think it necessary to foster the development of self-control.

An important question concerns whether these behaviors, which were showed by parents with a limited-malleable theory, are effective in promoting improved self-control in children. Specifically, it is not clear whether restricting children’s food and toy-related choices, promotes the development of self-control. Could restricting a child’s freedom give the child fewer opportunities to learn to restrain him or herself and, hence, rather undermine the development of self-control?

An additional question involves the transmission of beliefs about willpower from parents to children, and whether specific kinds of acts or ways of talking about willpower in parents foster harmful beliefs in children about willpower (see Gunderson et al., 2013; Haimovitz & Dweck, 2016). It is also important to keep in mind that the nonlimited willpower theory helps people exert self-control especially in the face of challenge, and this predicts better interpersonal outcomes (e.g., Moffitt et al., 2011; Tangney, Baumeister, & Boone, 2004). Indeed, being able to control one’s impulses and regulate one’s emotions in the face of high demands may be particularly crucial in challenging parenting situations and therefore contribute to relationship quality and functional parenting (Deater-Deckard, 2014; Valiente, Lemery-Chalfant, & Reiser, 2007).

Future Directions: Exploring the Antecedents of Willpower Theories

So far, most research on implicit theories about willpower has focused on their behavioral consequences in the laboratory and the field and mechanisms that explain these effects. Yet, little is known about their cultural, social, and developmental antecedents. Where do willpower theories come from?

Social Learning

Previous theoretical and empirical work suggests that one factor that shapes children's motivational beliefs, expectations, and values are their parent's beliefs (Eccles, 1993; Haimovitz & Dweck, 2016; Simpkins, Fredricks, & Eccles, 2012). For example, Simpkins et al. (2012) showed that when parents value a certain domain, like sports or literature, their children are more likely to develop an interest in that domain, too. How do parents' beliefs about willpower affect their children's beliefs?

According to social learning theory (Bandura, 1971; Bandura & Walters, 1963; Olson & Dweck, 2008), learning is a cognitive process that is tied to the social context of a person. One crucial element of this process is the observation of other people's behavior. Accordingly, children acquire knowledge about social norms by observing what their parents and other adults do, when they do it, and what consequences arise from this behavior. With regard to children's beliefs about willpower, we would assume that being raised by a person with a limited theory about willpower exposes a child to numerous adult behaviors implying that the capacity to exert effort is limited and that periods of hard work have to be followed by rest and recovery. As described above, we have found that people with a limited-resource theory strive for rest and recovery once they have exerted self-control (Job, Bernecker, Miketta, & Friese, 2015). Do parents with a limited theory rest more after they have exerted themselves as compared to parents with a nonlimited theory? Do they talk more about the need to rest after they have worked hard?

Previous research further shows that parents communicate their theories of intelligence in what they say to a child and by praise and feedback they provide to a child's performance (Gunderson et al., 2013; Haimovitz & Dweck, 2016; Mueller & Dweck, 1998). Do parents with a limited theory communicate to their child, explicitly or implicitly, that their child needs to rest after having worked hard ("You deserve a break!")? If so, do children of parents who have a limited theory infer that one has to rest after (demanding) work before being able to function well again?

Cultural Background

A recent series of studies (Savani & Job, in press) explored cross-cultural differences in willpower theories, which might inform their cultural roots. We found that in the United States, people tend to endorse a limited theory about willpower. Interestingly, in India, a country with a strong self-control tradition, we found the opposite—people tended to believe that completing strenuous mental tasks is energizing. Moreover, Indians exhibit a reversed ego depletion effect. They performed better after an initial demanding task, especially if they endorsed a non-limited theory of willpower—that is, who believed that exerting self-control is energizing.

These cultural differences could have their roots in philosophical traditions. Numerous religious traditions originating in India, including Hinduism, Jainism, and Buddhism, advocate the frequent exertion of self-control not just for monks and nuns but for lay people in their daily lives (Bronkhorst, 1993; Mosher, 2005; Walsh & Shapiro, 2006). Characteristics of the Indian schooling system as compared to the US might further contribute to the cross-cultural differences in theories about willpower. The workload of students in India is considerably higher than in the US putting less emphasis on breaks and times for recovery (Larson & Verma, 1999; Verma, Sharma, & Larson, 2002). Such a practice communicates to students that sustained mental effort is possible.

Future research should systematically investigate cultural differences in willpower beliefs and the socio-cultural mechanisms that perpetuate them. An interdisciplinary approach, including sociological and/or historical perspectives may generate knowledge of both theoretical and practical relevance. Indeed, research on the origins of willpower beliefs could inform, for instance, educational reforms and social policies on how to promote the development of nonlimited willpower theories in children and adults.

Interventions

Although a limited-resource theory might be functional in some situations, the accumulated evidence documents its costs. When people face high self-control demands those with a limited theory show impaired self-regulation, goal attainment, and well-being. An important direction for future research is to develop interventions that can help people adopt a nonlimited theory about willpower and self-regulate more effectively when they face high demands, such as in challenging academic programs or when a chronic disease requires a careful lifestyle change. Indeed, how to improve people's self-regulatory outcomes is a pressing issue (Diamond, 2012; Duckworth, Grant, Loew, Oettingen, & Gollwitzer, 2011).

Previous field-experiments show that it is possible to change people's implicit theories about intelligence and personality in field settings, with beneficial consequences including for academic performance (Aronson, Fried, & Good, 2002;

Blackwell, Trzesniewski, & Dweck, 2007; Paunesku et al., 2015; Yeager et al., 2016) and social outcomes (e.g., Yeager, Trzesniewski, & Dweck, 2013). Such interventions give people information (e.g., scientific reports) about the nature of human qualities and help them internalize this information using powerful persuasive techniques, such as “saying-is-believing” exercises in which people advocate for the intervention message to others (see Aronson, 1999; Yeager & Walton, 2011). Could this approach change people’s beliefs about willpower?

An important caution is that it would not be fruitful for people to infer that self-control is easy—that they have ample resources to resist temptation, say, and thus need not take normal steps to make self-regulation easier (e.g., putting the cookies on a high shelf). Ironically, simply learning that willpower is stronger than one might have supposed could backfire. People could feel encouraged to put themselves in situations they are ill-equipped to deal with (e.g., keeping temptations close at hand in the belief that they will be able to resist them indefinitely).

In a currently ongoing project, we have started to develop such an intervention. In a first study, participants did not learn that willpower is ample. Instead, materials emphasized that how you think about willpower matters, and you can choose how you think about it. They then thought of a person who struggles with willpower, and wrote a letter of advice to this person describing these ideas. In a first randomized field experiment with students enrolled in their first year at the university, we found that, for students who faced high academic demands, the intervention improved their academic self-regulation (i.e., time spent on academic tasks) and semester grades (Job, Flückiger, Bernecker, Lieb, & Mata, 2017). This gain was found relative to a control group exposed to parallel but psychologically neutral (in terms of theories about willpower) material that addressed time management. Thus, when confronted with high demands, students in the nonlimited willpower theory condition were able to scale up their academic effort to achieve greater success. Although these results are promising, many questions remain. Can such an intervention produce long-term change in people’s willpower theories and self-regulatory success? If so, what recursive processes contribute to lasting change? Can such an intervention be scaled-up to benefit a wide population? Can it be adapted to help non-student populations that face specific challenges, such as people trying to make lifestyle changes to manage a chronic disease?

Conclusion

Research on lay theories on self-control suggests that one reason people may fail to control themselves or have trouble reaching their personal goals involves their beliefs about self-control resources, not a true lack of resources. This approach does not deny that a person is in part an energy-based system. Obviously, people need food to function well, they get tired, and they need sleep. But in the normal range of self-regulatory demands people face in everyday life there is not a narrow energy-based constraint on self-control capacity. However, in a social and cultural

context that promulgates the belief that willpower depends on a highly limited resource—including in the strength model of self-control itself—people can easily believe in such a constraint, and thus attend and respond to minor fluctuations in their available resources. This belief then itself limits people's willpower.

Although research on willpower theories began as an alternative view of ego depletion, it extends beyond a mere critique of the strength model. Not only does it suggest that ego depletion is not an inevitable state determined by basic physiological processes. It further informs our understanding of processes, including self-efficacy and rest-goal activation, which contribute to self-regulatory performance, personal goal-striving, and well-being when self-regulatory demands arise. Thus, it brings top-down processes back into focus in self-regulation research, and shows that seemingly fixed, physiological principles can, at least in part, be created and modulated by people's beliefs and expectations.

References

- Aronson, E. (1999). The power of self-persuasion. *American Psychologist*, *54*, 584–588.
- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*, 113–125.
- Bandura, A. (1971). *Social learning theory*. New York, NY: General Learning Press.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122–147.
- Bandura, A., & Walters, R. H. (1963). *Social learning and personality development* (Vol. 14). New York, NY: Holt, Rinehart and Winston.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252–1265.
- Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*, 1–15.
- Bernecker, K., Hermann, M., Brandstätter, V., & Job, V. (2017). Implicit theories about willpower predict subjective well-being. *Journal of Personality*, *85*, 136–150.
- Bernecker, K., & Job, V. (2015a). Beliefs about willpower are related to therapy adherence and psychological adjustment in patients with type 2 diabetes. *Basic and Applied Social Psychology*, *37*, 188–195.
- Bernecker, K., & Job, V. (2015b). Beliefs about willpower moderate the effect of previous day demands on next day's expectations and effective goal striving. *Frontiers in Psychology*, *6*, 1496.
- Bernecker, K., & Job, V. (in press). Implicit theories about willpower in resisting temptations and emotion control. *Journal of Psychology*.
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263.
- Boule, N. G., Haddad, E., Kenny, G. P., Wells, G. A., & Sigal, R. J. (2001). Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: A meta-analysis of controlled clinical trials. *Journal of the American Medical Association*, *286*, 1218–1227.
- Brand-Miller, J., Petocz, P., Hayne, S., & Colagiuri, S. (2003). Low-glycemic index diets in the management of diabetes: A meta-analysis of randomized controlled trials. *Diabetes Care*, *26*, 2261–2267.

- Bronkhorst, J. (1993). *The two traditions of meditation in ancient India*. Delhi: Motilal Banarsidass.
- Brown, C. M., & McConnell, A. R. (2011). Discrepancy-based and anticipated emotions in behavioral self-regulation. *Emotion, 11*, 1091–1095.
- Brunstein, J. C. (1993). Personal goals and subjective well-being: A longitudinal study. *Journal of Personality and Social Psychology, 65*, 1061–1070.
- Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General, 144*, 796–815.
- Chambers, E. S., Bridge, M. W., & Jones, D. A. (2009). Carbohydrate sensing in the human mouth: Effects on exercise performance and brain activity. *The Journal of Physiology, 587*, 1779–1794.
- Chow, J. T., Hui, C. M., & Lau, S. (2015). A depleted mind feels inefficacious: Ego-depletion reduces self-efficacy to exert further self-control. *European Journal of Social Psychology, 45*, 754–768.
- Clarkson, J. J., Hirt, E. R., Jia, L., & Alexander, M. B. (2010). When perception is more than reality: The effects of perceived versus actual resource depletion on self-regulatory behavior. *Journal of Personality and Social Psychology, 98*, 29–46.
- Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The Mentor's Dilemma: Providing Critical Feedback Across the Racial Divide. *Personality and Social Psychology Bulletin, 25*, 1302–1318.
- Deater-Deckard, K. (2014). Family matters: Intergenerational and interpersonal processes of executive function and attentive behavior. *Current Directions in Psychological Science, 23*, 230–236.
- DeWall, C. N., Baumeister, R. F., Gailliot, M. T., & Maner, J. K. (2008). Depletion makes the heart grow less helpful: Helping as a function of self-regulatory energy and genetic relatedness. *Personality and Social Psychology Bulletin, 34*, 1653–1662.
- Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science, 21*, 335–341.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302.
- Duckworth, A. L., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. M. (2011). Self-regulation strategies improve self-discipline in adolescents: Benefits of mental contrasting and implementation intentions. *Educational Psychology, 31*, 17–26.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Taylor & Francis.
- Eccles, J. S. (1993). School and family effects on the ontogeny of children's interests, self-perceptions, and activity choices. In R. Dienstbier (Series Ed.) & J. E. Jacobs (Vol. Ed.), *Nebraska Symposium on Motivation: 1992. Developmental perspectives on motivation* (Vol. 40, pp. 145–208). Lincoln: University of Nebraska Press.
- Emmons, R. A. (1986). Personal strivings: An approach to personality and subjective well-being. *Journal of Personality and Social Psychology, 51*, 1058–1068.
- Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: The effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology, 87*, 557–572.
- Finkel, E. J., DeWall, C. N., Slotter, E. B., Oaten, M., & Foshee, V. A. (2009). Self-regulatory failure and intimate partner violence perpetration. *Journal of Personality and Social Psychology, 97*, 483–499.
- Fischer, P., Greitemeyer, T., & Frey, D. (2008). Self-regulation and selective exposure: The impact of depleted self-regulation resources on confirmatory information processing. *Journal of Personality and Social Psychology, 94*, 382–395.
- Fishbach, A., Shah, J. Y., & Kruglanski, A. W. (2004). Emotional transfer in goal systems. *Journal of Experimental Social Psychology, 40*, 723–738.

- Freeman, N., Shmueli, D., & Muraven, M. (2013). Lay theories of self-control influence judgments of individuals who have failed at self-control. *Journal of Applied Social Psychology, 43*, 1418–1427.
- Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review, 11*, 303–327.
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., et al. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology, 92*, 325–336.
- Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.
- Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1-3 year-olds predicts children's motivational frameworks 5 years later. *Child Development, 84*, 1526–1541.
- Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin, 136*, 495–525.
- Hagger, M. S., & Chatzisarantis, N. L. D. (2013). The sweet taste of success: The presence of glucose in the oral cavity moderates the depletion of self-control resources. *Personality and Social Psychology Bulletin, 39*, 28–42.
- Haimovitz, K., & Dweck, C. S. (2016). What predicts children's fixed and growth intelligence mind-sets? Not their parents' views of intelligence but their parents' views of failure. *Psychological Science, 27*, 859–869.
- Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science, 7*, 450–463.
- Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences, 18*, 127–133.
- Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science, 17*, 262–269.
- Job, V., Bernecker, K., Miketta, S., & Friese, M. (2015). Implicit theories about willpower predict the activation of a rest goal following self-control exertion. *Journal of Personality and Social Psychology, 109*, 694–706.
- Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion—Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science, 21*, 1686–1693.
- Job, V., Flückiger, L., Bernecker, K., Lieb, R., & Mata, J. (2017). An implicit theories about willpower intervention helps students to deal with demands: Effects on academic self-regulation and grades. *Manuscript in preparation*.
- Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2013). Beliefs about willpower determine the impact of glucose on self-control. *Proceedings of the National Academy of Sciences, 110*, 14837–14842.
- Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2015). Implicit theories about willpower predict self-regulation and grades in everyday life. *Journal of Personality and Social Psychology, 108*, 637–647.
- Klinger, E. (1977). *Meaning and void: Inner experience and the incentive in people's lives*. Minneapolis, MN: University of Minnesota Press.
- Kringelbach, M. L. (2004). Food for thought: Hedonic experience beyond homeostasis in the human brain. *Neuroscience, 126*, 807–819.
- Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evolutionary Psychology, 8*, 244–259.
- Larson, R. W., & Verma, S. (1999). How children and adolescents spend time across the world: Work, play, and developmental opportunities. *Psychological Bulletin, 125*, 701–736.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75–105). Hillsdale, NJ: Erlbaum.
- Little, B. (1983). Personal projects: A rationale and method for investigation. *Environment and Behavior, 15*, 273–309.

- Loschelder, D. D., & Friese, M. (2016). Moderators of the ego depletion effect. In E. R. Hirt, J. J. Clarkson, & L. Jia (Eds.), *Self-regulation and ego control* (pp. 21–42). San Diego, CA: Elsevier.
- Maier, G. W., & Brunstein, J. C. (2001). The role of personal work goals in newcomers' job satisfaction and organizational commitment: A longitudinal analysis. *Journal of Applied Psychology, 86*, 1034–1042.
- Masicampo, E. J., Martin, S. R., & Anderson, R. A. (2014). Understanding and overcoming self-control depletion. *Social and Personality Psychology Compass, 8*, 638–649.
- Martijn, C., Tenbült, P., Merckelbach, H., Dreezens, E., & De Vries, N. K. (2002). Getting a grip on ourselves: Challenging expectancies about loss of energy after self-control. *Social Cognition, 20*, 441–460.
- Miller, E. M., Walton, G. M., Dweck, C. S., Job, V., Trzesniewski, K. H., & McClure, S. M. (2012). Theories of willpower affect sustained learning. *PLoS ONE, 7*, e38680.
- Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science, 244*, 933–938.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., & Harrington, H. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*, 2693–2698.
- Molden, D. C., Hui, C. M., Scholer, A. A., Meier, B. P., Noreen, E. E., D'Agostino, P. R., et al. (2012). Motivational versus metabolic effects of carbohydrates on self-control. *Psychological Science, 23*, 1137–1144.
- Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin, 32*, 1024–1036.
- Mosher, L. (2005). *Praying: The Rituals of Faith* (Vol. 2). New York, NY: Seabury Books.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*, 33–52.
- Mukhopadhyay, A., & Johar, G. V. (2005). Where there is a will, is there a way? Effects of lay theories of self-control on setting and keeping resolutions. *Journal of Consumer Research, 31*, 779–786.
- Mukhopadhyay, A., & Yeung, C. W. (2010). Building character: Effects of lay theories of self-control on the selection of products for children. *Journal of Marketing Research, 47*, 240–250.
- Muraven, M., Shmueli, D., & Burkley, E. (2006). Conserving self-control strength. *Journal of Personality and Social Psychology, 91*, 524–537.
- Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin, 29*, 894–906.
- Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology, 74*, 774–789.
- Olson, K. R., & Dweck, C. S. (2008). A blueprint for social cognitive development. *Perspectives on Psychological Science, 3*, 193–202.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*, 784–793.
- Salmon, S. J., Adriaanse, M. A., De Vet, E., Fennis, B. M., & De Ridder, D. T. D. (2014). “When the going gets tough, who keeps going?” Depletion sensitivity moderates the ego-depletion effect. *Frontiers in Psychology, 5*, 647.
- Sanders, M. A., Shirk, S. D., Burgin, C. J., & Martin, L. L. (2012). The gargle effect: Rinsing the mouth with glucose enhances self-control. *Psychological Science, 23*, 1470–1472.
- Savani, K., & Job, V. (in press). Is ego-depletion a cultural phenomenon? Acts of self-control improve subsequent performance in cultures in which willpower exertion is believed to be energizing. *Journal of Personality and Social Psychology*.
- Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology, 85*, 33–64.

- Silvia, P. J., & Gendolla, G. H. E. (2001). On introspection and self-perception: Does self-focused attention enable accurate self-knowledge? *Review of General Psychology, 5*, 241–269.
- Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology, 48*, 1019–1032.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*, 271–324.
- Tice, D., Baumeister, R., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology, 43*, 379–384.
- Valiente, C., Lemery-Chalfant, K., & Reiser, M. (2007). Pathways to problem behaviors: Chaotic homes, parent and child effortful control, and parenting. *Social Development, 16*, 249–267.
- Verma, S., Sharma, D., & Larson, R. W. (2002). School stress in India: Effects on time and daily emotions. *International Journal of Behavioral Development, 26*, 500–508.
- Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology, 88*, 632–657.
- Vohs, K. D., Baumeister, R. F., & Schmeichel, B. J. (2012). Motivation, personal beliefs, and limited resources all contribute to self-control. *Journal of Experimental Social Psychology, 48*, 943–947.
- Vohs, K. D., & Faber, R. J. (2007). Spent resources: Self-regulatory resources availability affects impulse buying. *Journal of Consumer Research, 33*, 537–547.
- Walsh, R., & Shapiro, S. L. (2006). The meeting of meditative disciplines and Western psychology: A mutually enriching dialogue. *American Psychologist, 61*, 227–239.
- Webb, T. L., & Sheeran, P. (2003). Can implementation intentions help to overcome ego-depletion? *Journal of Experimental Social Psychology, 53*, 5–13.
- Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology, 55*, 493–518.
- Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., et al. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*, 374–391.
- Yeager, D. S., Trzesniewski, K. H., & Dweck, C. S. (2013). An implicit theories of personality intervention reduces adolescent aggression in response to victimization and exclusion. *Child Development, 84*, 970–988.
- Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267–301.

What Are People's Lay Theories About Mind Wandering and How Do Those Beliefs Affect Them?

Claire M. Zedelius and Jonathan W. Schooler

Many of the thoughts that pass through our minds each day are disconnected from the here and now. While we may seem engaged with our current activity or environment—our eyes scanning a page of text or locking with those of a conversation partner—our attention is often directed inwardly, to thoughts about current concerns, future plans, or fantasies. These types of thoughts, studied under the (somewhat interchangeably used) terms *stimulus-* or *task-unrelated thoughts*, *decoupled thought*, *daydreaming*, and *mind wandering* (e.g., Antrobus, 1968; Singer & Schonbar, 1961; Smallwood & Schooler, 2006), are often spontaneous and unsolicited (Seli, Carriere, & Smilek, 2014; Seli, Risko, Smilek, & Schacter, 2016), yet occupy an astonishing 30–50% of our waking life (Kane et al., 2007; Killingsworth & Gilbert, 2010; Klinger & Cox, 1987; McVay, Kane, & Kwapil, 2009), with far-reaching and often negative consequences for our performance (see Mooneyham & Schooler, 2013 for a review), mood (Franklin et al., 2013; Killingsworth & Gilbert, 2010), and safety (e.g., Galera et al., 2012). Sometimes stimulus-unrelated thoughts have an intrusive character. *Intrusive thoughts* tend to be repetitive and revolve around fears or traumatic events (Clark & Rhyno, 2005). Like normal mind wandering, intrusive thoughts are highly common among healthy individuals (Clark & Rhyno, 2005), but they are also a hallmark feature of a surprising range of disorders, including depression, generalized anxiety disorder, insomnia, obsessive–compulsive disorder (OCD), and posttraumatic stress disorder (PTSD; Clark, 2005; Davies & Clark, 1998). Given the ubiquity and impact of stimulus-unrelated thoughts among healthy individuals, and the role the intrusive variant of such thoughts seems to play in a variety of disorders, an important question, which we will examine in this chapter, is: how do ordinary people make

C.M. Zedelius (✉) · J.W. Schooler
Department of Psychological and Brain Sciences, University of California,
Santa Barbara, CA 93106, USA
e-mail: claire.zedelius@psych.ucsb.edu

sense of such thoughts? That is, what do we know about the character and the consequences different *lay theories* about stimulus-unrelated thoughts?

Why Lay Theories Matter

Lay theories are beliefs (also referred to as *naïve theories*, *implicit theories*, *folk theories*, *meta-cognitive beliefs*, or *mindsets*) that function as a “lens” through which people interpret events and make sense of their own and other people’s behavior (Dweck, Chiu, & Hong, 1995). Lay theories also inform people’s predictions about the consequences of their own behavior and the kind of behavior that can be expected from other people. What distinguishes lay theories from scientific theories is that lay theories are not necessarily explicit, and typically not rigorously formulated (Heider, 1958). Nonetheless, when asked, people usually have no difficulty reporting on their lay theories (see Plaks, Levy, & Dweck, 2009). What lay theories have in common with scientific theories is that they often divide people into different camps. And which “camp” somebody is in can influence the person’s behavior in important ways (e.g., see Burnette, O’Boyle, VanEpps, Pollack, & Finkel, 2013, for a meta-analytic review).

Consider the following example from research on lay theories in the domain of willpower: A popular lay theory about willpower (e.g., Job, Dweck, & Walton, 2010; Job, Walton, Bernecker, & Dweck, 2013, 2015; Martijn, Tenbült, Merckelbach, Dreezens, & de Vries, 2002), which—perhaps not coincidentally—is echoed in influential scientific theories (e.g., Baumeister, Vohs, & Tice, 2007; Gailliot et al., 2007; Muraven & Baumeister, 2000), holds that willpower relies on a finite biological resource that becomes “depleted” with use (and can be replenished by consuming glucose). Laypeople do not necessarily formulate the theory in quite these terms, but they may say that activities that demand a lot of willpower from them will leave them feeling “exhausted” or “drained” (and perhaps in need of a sugary snack), and that they are unable to exert any more willpower until they have rested or “refueled”. There is indeed evidence that exerting willpower causes a temporary state akin to depletion that can be reverted by glucose ingestion (see DeWall, Baumeister, Gailliot, & Maner, 2008; Gailliot & Baumeister, 2007; Gailliot et al., 2007; Gailliot, Peruche, Plant, & Baumeister, 2009; Hagger, Wood, Stiff, & Chatzisarantis, 2010). However, it has been found that this is true only for individuals who believe in the limited resource-theory. Those who do not believe in the theory show no depletion effect after exerting willpower (Carter, Kofler, Forster, & McCullough, 2015; Martijn et al., 2002; Job et al., 2010, 2013, 2015). Thus, the lay theory that willpower is a limited resource appears to affect people’s actual behavior by way of a self-fulfilling prophecy. This example powerfully illustrates how important it is that we make lay theories an integral part of our scientific theories. If we fail to take into account people’s beliefs about how they will think, feel, and behave in certain contexts, we can potentially derive incorrect or incomplete scientific theories about human cognition and behavior.

In the current chapter, we apply our focus on lay theories to the study of spontaneous stimulus-unrelated and intrusive thoughts. Over the last couple of decades, much progress has been made in gaining a better understanding of when people's minds wander off towards stimulus- or task-unrelated thoughts, what brain areas are involved in such mind wandering (e.g., Fox, Spreng, Ellamil, Andrews-Hanna, & Christoff, 2015), and what positive and negative consequences it has (e.g., Mooneyham & Schooler, 2013; Schooler et al., 2014). Only very recently have studies started to examine what kinds of beliefs laypeople have about this common experience, and how these beliefs may affect their tendency to mind wander. With regard to intrusive thoughts, there is a somewhat longer tradition of research focusing specifically on lay theories. This chapter brings together these different lines of research and gives an overview over their current state. First, we will discuss what kinds of—accurate or biased—lay theories people have about the frequency and meaning of spontaneous stimulus-unrelated thoughts. Next, we will focus on a key dimension that appears to be central to many lay theories: controllability. People are highly attuned to distinguishing between events and abilities that are under their personal control and events that lie outside of one's control (D'Andrade, 1987; Heider, 1958; Malle & Knobe, 1997; Molden & Dweck, 2006; Plaks, Grant, & Dweck, 2005). Recent evidence, which we will review in detail below, suggests that the extent to which people believe they have control over their tendency to mind wander affects how much they actually do so—similar to the self-fulfilling prophecy effect observed in the domain of willpower. Drawing on clinical psychology literature, we will also explore the impact of lay theories in the context of intrusive thoughts. There is extensive evidence that people's beliefs about the meaning and consequences of intrusive thoughts affect how people react to the occurrence of these thoughts. Particularly beliefs about the necessity to exert control over intrusive thoughts appear to play a key role in the development and maintenance of clinical disorders involving intrusive thoughts. Finally, we will discuss some future directions, focusing particularly on ways in which it may be possible to challenge or change people's theories about stimulus-unrelated thoughts.

How Much Mind Wandering Is Normal?

It is safe to say that most people mind wander a lot. Experience-sampling studies, in which participants are probed (e.g., via smart phones) at pseudo-random moments during their normal everyday activities, have consistently found that people mind wander circa 40% of the time (Cameron & Giuntoli, 1972; Franklin et al., 2013; Killingsworth & Gilbert, 2010; Klinger & Cox, 1987; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012). This figure is consistent with mind-wandering rates found during live or recorded lectures (Risko et al., 2012; Schacter & Szpunar, 2015; Wammes, Seli, Allan, Boucher, & Smilek, 2016), and during laboratory

experiments, in which participants are probed while performing reading, vigilance, or working memory tasks (e.g., Giambra, 1989, 1995; Grodsky & Giambra, 1990; McVay & Kane, 2009, 2010; McVay et al., 2009).

Procedures for assessing mind wandering vary between studies. Most typically, participants are probed at different times during their current activity and simply asked whether or not they were mind wandering just prior to being probed. Sometimes, participants are also asked to self-initiate a report whenever they notice that their mind has wandered off task. Of course, they can only “catch” these episodes themselves when they are aware that their thoughts have drifted away from the here and now—this is called mind wandering with *meta-awareness*. Research has shown that people often lack meta-awareness, however. That is, they fail to self-catch, but nonetheless report that they have been mind wandering when being probed by the computer, suggesting that they were not aware of the fact that they were mind wandering *until* being probed (Schooler, 2002; Schooler, Reichle, & Halpern, 2004; Schooler et al., 2011; Smallwood, McSpadden, & Schooler, 2007, 2008; Zedelius, Broadway, & Schooler, 2015).

Given that people often mind wander without meta-awareness, we were interested in how much people *think* they mind wander during day-to-day activities. Since thoughts are only ever directly accessible to the one having them, and we never know what goes on in other people’s heads, we also wanted to know whether people would underestimate (or overestimate) how much other people mind wander compared to them. To answer these questions, we collected data from a stratified sample of 1326 US Americans (Zedelius, Protzko, Schooler, 2017a). We asked them to estimate how much time they spend mind wandering during a normal day, and how much they thought other people mind wander on average, with the order of these questions counter-balanced. Finally, we also asked them whether they thought that they mind wander less, more, or about the same amount as other people. We found that people collectively are surprisingly accurate in their beliefs about what constitutes typical mind wandering rates. They estimated that people mind wander on average roughly 38% of the time. However, interestingly, we found that people tend to believe that they themselves mind wander somewhat less (roughly 33%) compared to others. This difference also emerged when participants were explicitly asked to make a comparative judgment about their own versus other people’s propensity to mind wander. Thus, despite estimating a rather typical average mind-wandering rate, they seem to display a self-serving bias when it comes to assessing their own mind wandering. This is in line with research showing that people generally tend to view themselves in a favorable light (Chambers & Windschitl, 2004; Dunning, Heath, & Suls, 2004; Pronin, Gilovich, & Ross, 2004; Williams & Gilovich, 2008). While mind wandering is not by definition negative or undesirable, very frequent mind wandering could be interpreted as an uncomfortable lack of control over one’s own thoughts. That said, it is possible that, if people were asked to report how much they themselves and the people around them had been mind wandering during a specific activity (say a lecture or a conversation), people may be more accurate at judging their own mind wandering and underestimate how much other people mind wander, simply because it is not easy to detect

mind wandering in other people. Thus, the self-serving bias may disappear when people reflect on their mind wandering during a particular situation rather than their general tendency to mind wander, a possibility that should be further investigated.

Lay Theories About the Meaning of Spontaneous Thoughts

According to experience-sampling research, the majority of spontaneous task-unrelated thoughts revolve around current concerns and plans for the future (e.g., Baird, Smallwood, & Schooler, 2011; Bernsten & Jacobsen, 2007; D'Argembeau, Renaud, & Van der Linden, 2011; Klinger, 2009, 2013; Klinger & Cox, 1987; Smallwood, Nind, & O'Connor, 2009). Thus, it seems fair to say that most stimulus-unrelated thoughts are relatively mundane in content, and not too dissimilar from more deliberate and task-related thoughts. The focus on current concerns and future plans also suggests that mind wandering may be functional, in that it can help with autobiographical planning and everyday problem solving (Baird et al., 2011; Bernsten & Jacobsen, 2007; Morsella, Ben-Zeev, Lanska, & Bargh, 2010). There is further evidence that mind wandering can be helpful for spurring creative ideas and insights. For instance, it has been shown that, when people were working on a creative idea generation task, engaging in a period of mind wandering (compared to performing a demanding task that left little room for mind wandering) helped them come up with a greater number of creative ideas afterwards (Baird et al., 2012). Moreover, a greater self-reported tendency to mind wander during everyday activities has been found to be associated with more creative insights in a creative problem-solving task (Baird et al., 2012; Zedelius & Schooler, 2015). Thus, all in all, there is evidence that spontaneous stimulus-unrelated thoughts, while often trivial in content, can be useful for planning, problem solving, and creative insights.

Research on lay theories of mind wandering has not directly addressed people's beliefs about the function of spontaneous stimulus-unrelated thoughts for *creative* insights per se, but suggests that common lay theories attribute a somewhat similar special meaning to spontaneous stimulus-unrelated thoughts. It has long been thought that semi-spontaneous thoughts—elicited through methods such as hypnosis, free association, or projective tests—reveal important insights, typically concerning an individual's secret motives and desires (Cramer, 1991; Holmes, 1968; Jacoby & Kelley, 1992; Murray, 1951; Poole, Lindsay, Memon, & Bull, 1995; Wegner & Smart, 1997). Inspired by this idea, Morewedge and colleagues reasoned that stimulus-unrelated thoughts that occur to us entirely spontaneously and during everyday activities could be interpreted by laypeople as providing meaningful self-insights due to our perceived lack of control over these thoughts (Morewedge, Giblin, & Norton, 2014). If a thought occurs to us uncontrollably and for no apparent reason, so the supposed lay theory goes, the thought must be personally meaningful. The authors indeed found that participants judged various types of spontaneous stimulus-unrelated thoughts, including dreaming, mind

wandering, and Freudian slips, as more meaningful and providing more self-insight than more deliberate thoughts. Moreover, participants rated the *same* thought (e.g., a positive or negative childhood memory) as more insightful when they imagined it occurring to them spontaneously during an unrelated task, then when they were simply asked to try to recall it. Thus, lay theories about the meaning of spontaneous stimulus-unrelated thoughts show some overlap with scientific theories in that both consider stimulus-unrelated thoughts a potential source of valuable insights. However, considering that most stimulus-unrelated thoughts revolve around mundane content, the special meaning laypeople attribute to those thoughts seems exaggerated.

Lay Theories About the Controllability of Stimulus-Unrelated Thoughts

We have speculated that people's lay theories about how much they mind wander and what meaning it has are pervaded by an experienced lack of control over their thoughts. Indeed, the fact that we spend a substantial amount of time engaged in stimulus-unrelated thoughts, often without being aware of it (Schooler et al., 2011; Seli et al., 2016) and despite the numerous costs frequently associated with it (Mooneyham & Schooler, 2013), suggests that we lack control over our stimulus-unrelated thoughts. On the other hand, there is some evidence that mind wandering is not entirely uncontrollable. First, people *sometimes* mind wander intentionally, for instance to cope with boredom (Seli et al., 2016). Moreover, it has been found that people mind wander more at "opportune" moments, that is, when task demands are low, then when a task requires their full attention (e.g., Levinson, Smallwood, & Davidson, 2012; Smallwood & Schooler, 2006; Smallwood, Obonsawin, & Reid, 2003). This suggests that people exert at least some amount of control over when they let their minds wander. One could counterargue that this type of context-dependent mind wandering does not reflect active, strategic control, but that people simply *cannot* mind wander as much when their working memory resources are tied up by a demanding task (Smallwood, 2010; see also McVay & Kane, 2010). However, the tendency for "opportune" or "strategic" mind wandering (Franklin et al., 2014) is greater among individuals with larger working memory capacity (Kane et al., 2007; Rummel & Boywitt, 2014). Thus if it were the case that demanding tasks only reduce mind wandering because they tax-limited working memory resources, we would expect that high-capacity individuals mind wander more during demanding tasks. Instead, they mind wander less during these tasks (yet more during undemanding tasks). Finally, evidence that people have at least some, however limited, control over their spontaneous stimulus-unrelated thoughts comes from the finding that people self-catch more stimulus-unrelated thoughts when they are more motivated to do so (Zedelius, Broadway, & Schooler, 2015).

In light of the mixed evidence for controllability, we expected that people would have different lay theories about the topic. Some may view mind wandering as something that is largely outside their control, due perhaps to spontaneous attentional fluctuations inherent in the functioning of the brain. Others may interpret episodes of mind wandering as failure on their part to pay attention or control their thoughts. Such different beliefs seem to be reflected in the way people talk about mind wandering. Sometimes people talk about it in a way that suggests passivity and a lack of control, using phrases such as: “My mind has wandered off again,” or “My thoughts got carried away.” Sometimes people use more actively formulated phrases that imply a certain level of personal control and responsibility, such as “I wasn’t paying attention” or “I was thinking about something else”. Moreover, we routinely demand of children or students that they pay attention to their current tasks and activities. Based on the principle that “ought implies can”, demanding of others that they control their thoughts betrays the implicit theory that they *can* control their thoughts (Scruton, 1982; Stern, 2004, Vranas, 2007).

To formally assess people’s beliefs about their ability to control their wandering minds, and examine whether these beliefs affect actual mind-wandering rates by way of a self-fulfilling prophecy, we developed a novel scale that assesses the extent to which individuals agree or disagree with statements presenting mind wandering as something that is controllable or largely outside of personal control (e.g., “Even in moments when it really matters, I can’t do much to keep my mind from wandering”; “How much people mind wander is something about them that they can’t change very much”; Zedelius, Protzko, & Schooler, 2017b). In several studies, conducted online with a community sample and in the laboratory with students, we found that scores on the scale varied among individuals, confirming our expectation that individuals have different lay theories about mind wandering and controllability. More interestingly, we found that individual differences in these lay theories predicted participants’ self-reported mind-wandering rates during everyday activities (Study 1) as well as their probe-caught mind-wandering rates during a reading task in the laboratory, in which they read a short fictional text for comprehension (Studies 2–3). Individuals who believed that mind wandering is controllable reported fewer mind wandering episodes than those who believed that it is uncontrollable. Importantly, they also showed increased reading comprehension. Thus, it seems that people who believe that they have more control over their mind wandering actually mind wander less. An explanation for this finding is that people who believe that they have more control over their mind wandering regulate their attention more. This is in line with previous research showing that people are more likely to self-regulate their behavior when they believe that they have control over the behavior (e.g., Burnette et al., 2013).

The relationship between people’s lay theories about mind wandering and their actual tendencies to mind wander can go in both directions. Individuals who mind wander a lot may infer that mind wandering must be very difficult or impossible to control. To show that lay theories about mind wandering can also affect mind-wandering rates, we experimentally manipulated lay theories

(Zedelius et al., 2017b, Study 4) by presenting participants with instructions promoting either a “controllable” theory or an “uncontrollable” theory, or neutral instructions. In all three conditions, participants were given the same definition of mind wandering and were told that mind wandering often happens spontaneously. The only difference was that we told participants that people are “surprisingly good at controlling their mind wandering”, “simply by deciding to focus”, or that mind wandering is “very hard if not impossible to control”, and that people often fail to control their mind wandering “despite trying very hard to focus”. After receiving instructions, participants performed the same reading task used in the earlier correlational studies (reading for comprehension), during which they were intermittently probed and asked whether or not they had been mind wandering just prior to the probe. The results showed that participants in the controllable condition reported fewer probe-caught mind wandering episodes and higher reading comprehension scores than those in the uncontrollable condition.

More research is needed to test the mechanism behind this effect. The (chronic or experimentally induced) belief that mind wandering is controllable could facilitate sustained attention, reducing the occurrence of task-unrelated thoughts. Alternatively it is possible that the belief leads people to *notice* task-unrelated thoughts more, enabling them to redirect their attention to the task, and thus making mind-wandering episodes more short-lived and less disruptive. It is also possible that individuals who believe that mind wandering is controllable (vs. uncontrollable) are better at regulating their attention in line with current task demands, leading them to mind wander more at opportune moments and less when demands are high. Finally, it is in principle possible that lay theories about mind wandering affect people’s *interpretations* of what constitutes task-unrelated thought more than actual incidences of mind wandering, thus leading to biased self-reports. Note that this latter explanation, however, does not account for the finding that participants with a “controllable” lay theory of mind wandering reported not only fewer mind wandering episodes but also showed increased reading comprehension. That said, future research is necessary to examine these different potential mechanisms to help us get a better understanding of how lay theories affect people’s reported and actual mind wandering.

Another unanswered question is where lay theories about the controllability of spontaneous stimulus-unrelated thoughts originate. In one study, we found that lay theories about the controllability of mind wandering correlated, albeit weakly, with lay theories about intelligence (Zedelius et al., 2017b, Study 1). Individuals who believed more strongly that one’s tendency to mind-wander is something that is in principle controllable, were also slightly more likely to endorse the belief that intelligence is a skill that can be improved with practice. This suggests that these distinct lay theories may fit into a network of interrelated lay beliefs about cognitive capacities. More research is needed to investigate how different lay theories relate to each other and if they are causally linked. If so, it may be the case that strengthening one lay theory, say, that mind wandering is controllable, may also strengthen related beliefs, such as the belief that intelligence or other mental capacities can grow and be developed.

Lay Theories About Intrusive Thoughts

Beliefs about controllability have received particular attention in research on intrusive thoughts. Intrusive thoughts are spontaneous stimulus- or task-unrelated thoughts whose occurrence is experienced as unwanted or unacceptable (e.g., Rachman, 1981). Intrusive thoughts are not by definition negative in content. People also experience intrusive thoughts with positive or mixed content. Those can be thoughts that occur with a disruptive suddenness or intensity or thoughts that are deemed taboo, such as thoughts with sexual content or thoughts about a former romantic partner or a romantic alternative (Baird, Smallwood, Fishman, Mrazek, & Schooler, 2013; Byers, Purdon, & Clark, 1998; Bywaters, Andrade, & Turpin, 2004; Clark, Purdon, & Byers, 2000; Gonzaga, Haselton, Smurda, Davies, & Poore, 2008; Reynolds & Salkovskis, 1992). Most intrusive thoughts, however, revolve around worries and fears (Clark & de Silva, 1985). Moreover, following a minor or major traumatic event, it is common for individuals to reexperience the event in the form of intrusive thoughts and memories (Clark, 2005; Davies & Clark, 1998; Reynolds & Brewin, 1999).

There is considerable evidence that healthy individuals routinely experience intrusive thoughts (e.g., Bywaters et al., 2004; Clark, 2005; Rachman & de Silva, 1978; Sarason, Pierce, & Sarason, 1996; Wegner & Pennebaker, 1993), without necessarily experiencing great psychological distress, and sometimes even without being consciously aware of them (Baird et al., 2013). Yet, frequent intrusive thoughts can be a source of distress, and such thoughts are a defining feature of many clinical disorders; Depression and generalized anxiety disorder are associated with intrusive thoughts akin to rumination and self-doubt (Harrington & Blankenship, 2002; Reynolds & Brewin, 1999; Wenzlaff, Wegner, & Roper, 1988). Individuals suffering from insomnia report frequent intrusive thoughts related to their inability to sleep (e.g., Harvey, 2002; Harvey & Payne, 2002; Wicklow & Espie, 2000). OCD is characterized by repetitive intrusive thoughts and impulses, often involving some form of harm (Rachman, 1997), and PTSD is defined by frequent intrusive thoughts and memories related to a traumatic event (Clark, 2005; Davies & Clark, 1998; Reynolds & Brewin, 1999).

Interestingly, research has shown that the intrusive thoughts experienced by healthy individuals are surprisingly similar in form and content (albeit comparatively less frequent) to the intrusive thoughts observed in clinical populations (e.g., Clark & Rhyno, 2005; Rachman & de Silva, 1978; Radosky et al., 2014; Sarason et al., 1996; Wegner & Pennebaker, 1993). Yet, individuals with disorders involving intrusive thoughts, compared to healthy individuals, generally perceive these thoughts as considerably more distressing, anxiety provoking, and difficult to control (Rachman & de Silva, 1978). What, then, explains the markedly different responses of these populations to rather similar types of mental events? Research suggests that an important difference between the intrusive thoughts of healthy individuals and those symptomatic of a disorder lies in people's lay theories about

the meaning and consequences of intrusive thoughts, particularly those related to the perceived need to control one's thoughts.

In the context of obsessive–compulsive disorder, Rachman (1997) proposed that intrusive thoughts are more likely to cause distress and develop into obsessions if an individual believes them to reveal insights into their true motives and desires. As discussed before, people generally tend to attribute such meaning to spontaneous stimulus-unrelated thoughts (Morewedge et al., 2014). While this tendency may be relatively inconsequential when it comes to mundane mind wandering (e.g., about a recent memory or a future activity), it can have severe consequences in the case of intrusive thoughts, which are often perceived as immoral, disgusting, or dangerous. The thought of causing another person harm, for instance, can lead a person to believe that he or she is fundamentally evil or worthless, a belief that causes understandable distress. Other common lay theories that can exacerbate the distress are, first, the belief that merely *having* a particular thought is itself dysfunctional or immoral, and, second, that one is likely to act upon one's thoughts. This latter belief is also referred to as *thought-action fusion* (see Shafran & Rachman, 2004 for a review). Both these beliefs have been found to be common among patients with OCD and PTSD (e.g., Bryant & Guthrie, 2005; Owens, Chard, & Cox, 2008). Moreover, these beliefs often lead to yet another, related, belief; the belief that it is necessary to gain full control over one's intrusive thoughts (Obsessive Compulsive Cognitions Working Group, 2005; Shafran, Thordardson, & Rachman, 1996).

The most drastic form of exerting control over one's thoughts is thought suppression—blocking the very occurrence of unwanted thoughts from the stream of consciousness. Thought suppression is a strategy often adopted spontaneously in response to intrusive thoughts (e.g., Clark & Purdon, 2009), and also an approach sometimes used in therapeutic treatments of disorders involving intrusive thoughts. One of the oldest, and still widely used, treatments for obsessive–compulsive disorder, for instance, is thought stopping, in which patients are taught to respond to intrusive thoughts by saying or thinking the word “stop” (or sometimes performing a corresponding action such as snapping a rubber band on the wrist; Hannan & Tolin, 2005; Wolpe, 1990). There is evidence, however, that thought stopping or suppression is mostly ineffective. Attempts to suppress thoughts, while effective for a very short period of time, can lead to rebound effects; that is, an increase rather than decrease in the number of intrusive thoughts, over time (for a meta-analysis, see Abramowitz, Tolin, & Street, 2001; for more recent studies, see Iijima & Tanno, 2012; Koster, Rassin, Crombez, & Näring, 2003; Lambert, Hu, Magee, Beadel, & Teachman, 2014; Logel, Iserman, Davies, Quinn, & Spencer, 2009), along with increased psychological distress (Beck, Gudmundsdottir, Paylo, Miller, & Grant, 2006; Wegner & Gold, 1995). To make things worse, rebound effects following attempts to suppress thoughts have also been shown to diminish the belief that thoughts are controllable (Beck et al., 2006; Magee, Harden, & Teachman, 2012), and increase psychological distress and self-blaming (Clark & Purdon, 2009; Magee & Teachman, 2007).

Different explanations for rebound effects following thought suppression have been proposed. According to Wegner's ironic monitoring theory (Wegner, 1994),

attempts to suppress a particular thought engage two processes: For one, a conscious control or operating process deliberately diverts attention away from the unwanted thought by bringing to mind other, unrelated thought content, such as when one thinks of happy memories or lists items on one's mental grocery list to refute an unwanted or intrusive thought. At the same time, a monitoring process scans the stream of consciousness to detect unwanted thought content when it intrudes. Ironically, in order to detect an unwanted thought when it occurs, the monitoring process has to also keep the thought itself and related thought content accessible in mind. Due to this heightened accessibility, the thought will, from time to time, enter the stream of consciousness, and the more other topics have been exhausted, the more likely it becomes that one's mind returns to the suppressed thought (Wegner & Erber, 1992). Moreover, when attentional resources are diverted from the process of actively generating unrelated thought content, the unwanted thought becomes more likely to occur. This makes for an increased rebound effect under conditions of high cognitive load or after an extended period of thought suppression (Wegner & Erber, 1992; Wenzlaff & Wegner, 2000).

While Wegner's theory remains the most prominent explanation of rebound effects following thought suppression, another, not necessarily alternative but complementary, explanation has been put forward, which places an emphasis on the role of beliefs. According to the motivational inference model (Förster & Liberman, 2001, 2004), individuals interpret the effort involved in suppressing intrusive thoughts as diagnostic of how much they want to engage in those thoughts or in activities associated with the thoughts. For example, if a person finds it very difficult to suppress thoughts about a former romantic partner, they may conclude that they must have a strong desire to think of that person, or to reengage with them. This, in turn, may encourage the person to engage more, rather than less, in the intrusive thought. Förster and Liberman tested this model in a number of studies. In one study (Förster & Liberman, 2001, Study 1), they instructed participants to suppress thoughts of a white bear. In one condition (high motivation condition), they then experimentally induced the lay theory central to the motivational inference model in participants by simply telling participants that having difficulty suppressing thoughts of white bears indicates a high motivation to think of white bears. In a second condition (low motivation condition), they gave opposite instructions (i.e., having difficulty suppressing thoughts of white bears indicates low motivation to think these thoughts), and in a control condition, they did not induce a lay theory. While the conditions did not differ in their *initial* success at suppressing thoughts of white bears, participants in the high motivation condition indeed showed a greater rebound effect than participants in the low motivation condition. Moreover, participants in the control condition showed the same strong rebound effect as those in the high motivation condition. These results suggest that the rebound effect is indeed caused by spontaneous motivational inferences.

Thus, to sum up the above research, different lay theories can exacerbate intrusive thoughts, both in frequency and in the psychological distress they cause, in several ways: first, lay theories about the meaning and consequences of intrusive thoughts—specifically, the beliefs that unwanted thoughts convey insights into

one's motives, that having unwanted thoughts is dysfunctional or immoral, and that one is likely to act upon one's intrusive thoughts—can cause distress and motivate individuals to try to control or suppress their thoughts. Second, attempts to suppress intrusive thoughts can increase the frequency and intensity of these thoughts by way of a rebound effect. Third, the belief that failures at thought suppression convey motivation to engage with an unwanted thought can further increase the rebound effect.

Earlier, we discussed evidence suggesting that the belief that one has control over one's spontaneous stimulus-unrelated thoughts, in the context of normal mind wandering, is associated with a reduced number of stimulus-unrelated thoughts. How does this finding relate to the seemingly contradicting findings that attempts to control intrusive thoughts so often lead to rebound effects? First, we note the distinction between having control over one's general tendency to mind wander and exerting control over the occurrence of a specific unwanted thought. Although the exact mechanisms that allow us to control our tendency to mind wander are not clearly defined, it likely requires the capacity to maintain sustained attention and involvement in one's current task or environment and the capacity to refocus attention when noticing that the mind has wandered off. These processes are different from the suppression of a particular thought. When suppressing a particular thought, processes such as Wegner's ironic monitoring and motivational inferences are much more applicable. One can more easily monitor a particular thought than a range of all kinds of thoughts that have the potential to come up during some activity. And having difficulty suppressing a particular thought can more easily give rise to the inference that one must have a strong desire to engage with this thought. Still, these processes are likely not entirely absent in people's responses to normal mind wandering. For instance, observing that one's mind frequently wanders off to a particular topic could lead an individual to try to temporarily suppress thoughts about that topic, which could evoke ironic monitoring effects. And having difficulty preventing a range of different task-unrelated thoughts might not trigger motivational inferences about any particular thought, but it could lead a person to conclude that they would rather engage in anything else but their primary activity.

Perhaps a more important distinction relevant to explain the different effects of control beliefs on normal mind wandering and intrusive thoughts is between the belief that control over one's thoughts is *possible* and the belief that exerting complete control is *necessary*. For individuals who believe that their intrusive thoughts are immoral or unacceptable, or that merely thinking these thoughts makes it more likely that one will act on them, thought control is often seen as necessary (Purdon & Clark, 2002; Shafran, Thordarson, & Rachman, 1996). This belief has more serious implications than the belief that control is merely possible. If control is necessary, every intrusive or unwanted thought represents a failure. This could maintain the vicious circle of trying harder to control one's thoughts and thereby increasing the change for rebound effects. To successfully regulate one's thoughts, it may be critical to have an accepting attitude towards occasional failure. Thus, we

suggest that the belief that one has the potential to exert some control over one's spontaneous stimulus-unrelated thoughts can be beneficial, as long as one does not hold the dysfunctional belief that absolute control is necessary.

Concluding Remarks and Future Directions

Many of the thoughts that pass through our minds are spontaneous and entirely unrelated to our current task or environment. Given how frequent such thoughts are, and how surprising—and sometimes even distressing—they can be, it is intriguing to ask how people make sense of these experiences. What kinds of implicit or explicit theories do they have to explain why their mind wanders off to a memory of an old flame, why a student is not paying attention, or why a distressing unwanted thought keeps coming up in their mind? And how do these theories, in turn, affect their thoughts and behavior? While the last decades have seen much progress in understanding the causes and consequences of spontaneous stimulus-unrelated thoughts, it is only recently that we have begun to ask these questions about people's lay theories about spontaneous stimulus-unrelated thoughts. Most of this research has focused on theories revolving around meaning and controllability. As we have discussed in this chapter, people tend to ascribe important meaning to spontaneous stimulus-unrelated thoughts. Spontaneous thoughts are often believed to convey insights into our motives and desires. Sometimes, overstating the meaning of spontaneous thoughts can have negative consequences, as is the case for intrusive thoughts. Such thoughts are often evaluated on moral grounds and are sometimes believed to be precursors of harmful actions. People hold different beliefs about the degree to which spontaneous stimulus-unrelated thoughts are controllable or outside their control. We have discussed recent findings showing that these beliefs can affect how much people actually mind wander. Individuals who believe that they can control their spontaneous thoughts mind wander less. We have also discussed the role dysfunctional control-related beliefs play in mental disorders involving intrusive thoughts. Particularly the belief that it is necessary to fully control one's unwanted thoughts can increase the frequency and intensity of these thoughts.

The research discussed in this chapter illustrates the important role that control-related beliefs play in how much people experience spontaneous stimulus-unrelated thoughts and how they respond to such thoughts. Thus, an intriguing question for future research is how to induce adaptive changes in people's lay beliefs in order to help them minimize disruptive mind wandering and distress in response to intrusive thoughts.

One approach to challenging people's lay theories is to be fairly explicit, for instance by simply telling them that they have a good amount of control over how much they mind wander during lectures or in class. This approach has been effective in the short term in the laboratory (Förster & Liberman, 2001, 2004; Reuven-Magril, Rosenman, Liberman, & Dar, 2009; Zedelius et al., 2017b),

but it is an open question whether it is equally effective in real-life settings and whether it has long-lasting effects. Encouraging evidence that such an approach could work comes from successful real-life interventions targeting lay theories related to intelligence (Blackwell, Trzesniewski, & Dweck, 2007; Burnette & Finkel, 2012; Yeager & Dweck, 2012; Yeager & Walton, 2011). However, it needs to be explored whether it is effective for theories about spontaneous and intrusive thoughts. In this area, people may have strongly held preexisting beliefs that may be anchored in moral or religious belief systems (e.g., concerning the meaning and moral implications of taboo thoughts; Berman, Abramowitz, Pardue, & Wheaton, 2010), and are more resistant to change. Moreover, changing people's lay theories in this way may be particularly difficult when people's day-to-day experiences seem to provide contradicting feedback. The belief that one has control over one's wandering mind, for instance, is easily challenged by frequent and frustrating experiences of mind-wandering in everyday life, or by frequent experiences of unwanted and intense intrusive thoughts.

Achieving long-lasting, stable change may require an approach that contains both explicit instructions about the efficacy of mental control while also providing practical, training to maintain sustained attentiveness. A training that seems to practice this skill is mindfulness meditation. Mindfulness meditation is derived from the Eastern Vipassana meditation (Kabat-Zinn, 2003) and typically involves focusing attention on one object (e.g., one's breath) and returning to it after noticing that one's mind has wandered. Supporting the view that mindfulness and mind-wandering are opposing constructs (Mrazek, Smallwood, & Schooler, 2012), studies have shown that a brief mindfulness exercise can temporarily reduce mind-wandering (Mrazek et al., 2012; Zeidan, Johnson, Diamond, David, & Goolkasian, 2010), and that mindfulness-meditation training over several weeks yields further improvements in attentional control (Jensen, Vangkilde, Frokjaer, & Hasselbalch, 2012; Jha et al., 2015; Mrazek, Franklin, Phillips, Baird, & Schooler, 2013; Posner, Rothbart, & Tang, 2015; Sedlmeier et al., 2012; Tang & Posner, 2009). Observing such improvements in oneself may also strengthen the belief that one has control over one's thoughts, a belief that, in turn, could further improve the effectiveness of the training itself, thus producing a mutually reinforcing effect of instruction and training. There is indeed preliminary evidence that mindfulness training increases people's endorsement of the belief that they have control over their spontaneous stimulus-unrelated thoughts (Mrazek et al., 2017). To what extent this belief can further increase the effectiveness of the training is a question that needs to be examined in future research.

We should note that any approach to changing people's lay theories of mind wandering in order to help them maintain better focus or experience less distress in response to intrusive thoughts should focus on promoting the belief that control is possible to some extent, but that full control is not the goal. As we have discussed earlier in the context of intrusive thoughts, the belief that it is necessary to gain full control over one's unwanted thoughts is particularly likely to lead to rebound effects and increased distress. In most mindfulness instructions, it is emphasized that complete thought control is neither possible nor desirable. It is understood that

spontaneous and sometimes intrusive thoughts are a constant part of the stream of consciousness. Instead of trying to control or suppress thoughts, practitioners are instructed to observe their thoughts in a detached, nonreactive manner and with a nonjudgmental attitude. It is therefore expected that engaging in mindfulness-based training will increase the likelihood that a person responds to intrusive thoughts with acceptance rather than suppression, a response style that has been shown to be associated with reduced suppression-related rebound and reduced psychological distress (Marcks & Woods, 2005, 2007; Najmi, Riemann, & Wegner, 2009). Thus, by fostering an accepting attitude, mindfulness-based training may be a good approach to challenging the dysfunctional belief that it is necessary to gain complete control over the occurrence of unwanted intrusive thoughts.

In recent years, it has been proposed that disorders involving intrusive thoughts can be treated effectively by challenging dysfunctional beliefs about control through a combination of explicit instruction and more experiential approaches. Marcks and Woods (2005), for instance, have noted that “creating effective acceptance-based procedures can be a challenge, since acceptance is naturally counter-intuitive. Furthermore, acceptance cannot be manipulated through simple instructions (i.e., “accept thought x”), but rather it must be done experientially.” In cognitive or meta-cognitive therapies for OCD patients are often encouraged to seek out short exposure to stimuli or environments eliciting intrusive thoughts, a process that is thought to break thought-action fusion because the patient experiences directly that they do not, as is their fear, act on unwanted thoughts (e.g., Fisher & Wells, 2005; Wells, 2005). Mindfulness-based therapy has also been proposed to be a useful approach to treating OCD, again because adopting a detached perspective towards one's own thoughts is thought to reduce thought-action fusion and the belief that one must control unwanted intrusions (Hannan & Tolin, 2005; Wells, 2005; Wilkinson-Tough, Bocci, Thorne, & Herlihy, 2010; Wegner, 2011).

We have devoted much attention to the aspect of control in people's lay theories about spontaneous stimulus-unrelated thoughts. There are many other aspects to people's lay theories that future research could examine. For instance, we know from mind wandering research that engaging in stimulus-unrelated thoughts has widespread negative, and also some positive consequences. What beliefs do people have about how mind wandering affects them? Do these beliefs shape when or how much people mind wander, or how they experience it?

There is evidence that engaging in stimulus-unrelated thoughts often brings people in a more negative mood (Franklin et al., 2013; Killingsworth & Gilbert, 2010; although highly interesting mind wandering episodes can increase people's; see Franklin et al., 2013). Moreover, there is overwhelming evidence that mind wandering interferes with performance on even the simplest of tasks (see Mooneyham & Schooler, 2013). To what extent are people aware of these effects—both generally, and when it comes to *their own* tendency to mind wander? Do people believe that mind wandering makes them unhappy, or do they think it offers a pleasant distraction from boring activities? Do people hold different beliefs about how much mind wandering hurts their performance? If so, these beliefs could affect

how much people regulate their thoughts, similarly to what we have found for control-related beliefs. There is also evidence that engaging in stimulus-unrelated thoughts can be a source of creativity (Baird et al., 2012; Zedelius & Schooler, 2015). To what degree is this reflected in people's lay theories? Do some people embrace mind wandering episodes more than others because they believe it to bring about valuable creative insights? What other specific beliefs do people have about how spontaneous stimulus-unrelated thoughts influence their lives? These are all questions for future research that go beyond the question of how much control we have over our thoughts.

In conclusion, the present chapter has provided an overview of the relatively young field of research devoted to understanding how laypeople make sense of the spontaneous stimulus-unrelated thoughts that pass through their minds during much of their waking life. The research illustrates that people do not believe these thoughts to be merely trivial distractions of their day-to-day activities. It also shows that people's beliefs or lay theories matter. They affect how much people let their minds wander or regulate their thoughts, and how they respond to unwanted thought intrusions. Moreover, dysfunctional beliefs can lead people to engage in ineffective thought-control strategies, which cause much psychological distress. Many questions remain to be explored before we can truly appreciate the full manner in which people's lay theories about mind wandering affect their day-to-day lives, but it is clear that these questions deserve our attention.

References

- Abramowitz, J. S., Tolin, D. F., & Street, G. P. (2001). Paradoxical effects of thought suppression: A meta-analysis of controlled studies. *Clinical Psychology Review, 21*(5), 683–703.
- Antrobus, J. S. (1968). Information theory and stimulus-independent-thought. *British Journal of Psychology, 59*, 423–430.
- Baird, B., Smallwood, J., Fishman, D. J., Mrazek, M. D., & Schooler, J. W. (2013). Unnoticed intrusions: Dissociations of meta-consciousness in thought suppression. *Consciousness and Cognition, 22*(3), 1003–1012.
- Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction mind wandering facilitates creative incubation. *Psychological Science, 0956797612446024*.
- Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition, 20*(4), 1604–1611.
- Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16*, 351–355.
- Beck, J. G., Gudmundsdottir, B., Palyo, S. A., Miller, L. M., & Grant, D. M. (2006). Rebound effects following deliberate thought suppression: Does PTSD make a difference? *Behavior Therapy, 37*(2), 170–180.
- Berman, N. C., Abramowitz, J. S., Pardue, C. M., & Wheaton, M. G. (2010). The relationship between religion and thought–action fusion: Use of an in vivo paradigm. *Behaviour Research and Therapy, 48*(7), 670–674.
- Bernsten, A., & Jacobsen, A. S. (2007). Involuntary (spontaneous) mental time travel into the past and future. *Consciousness and Cognition, 17*, 1093–1144.

- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246–263. doi:10.1111/j.1467-8624.2007.00995.x
- Burnette, J. L., & Finkel, E. J. (2012). Buffering against weight gain following dieting setbacks: An implicit theory intervention. *Journal of Experimental Social Psychology, 48*, 721–725. doi:10.1016/j.jesp.2011.12.020
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139*(3), 655.
- Byers, E. S., Purdon, C., & Clark, D. A. (1998). Sexual intrusive thoughts of college students. *Journal of Sex Research, 35*(4), 359–369.
- Bywaters, M., Andrade, J., & Turpin, G. (2004). Intrusive and non-intrusive memories in a non-clinical sample: The effects of mood and affect on imagery vividness. *Memory, 12*(4), 467–478.
- Bryant, R. A., & Guthrie, R. M. (2005). Maladaptive appraisals as a risk factor for posttraumatic stress. *Psychological Science, 16*, 749–752. doi:10.1111/j.1467-9280.2005.01608
- Cameron, P., & Giuntoli, D. (1972). Consciousness sampling in the college classroom or is anybody listening? *Intellect, 101*, 63–64.
- Carter, E., Kofler, L., Forster, D., & McCullough, M. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*. doi:10.1037/xge0000083
- 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.
- Clark, D. A. (2005). *Intrusive thoughts in clinical disorders: Theory, research, and treatment*. New York: Guilford Press.
- Clark, D. A., & de Silva, P. (1985). The nature of depressive and anxious, intrusive thoughts: Distinct or uniform phenomena? *Behaviour Research and Therapy, 23*(4), 383–393.
- Clark, D. A., & Purdon, C. (2009). Mental control of unwanted intrusive thoughts: A phenomenological study of nonclinical individuals. *International Journal of Cognitive Therapy, 2*(3), 267–281.
- Clark, D. A., Purdon, C., & Byers, E. S. (2000). Appraisal and control of sexual and non-sexual intrusive thoughts in university students. *Behaviour Research and Therapy, 38*(5), 439–455.
- Clark, D. A., & Rhyno, S. (2005). Unwanted intrusive thoughts in nonclinical individuals: Implications for clinical disorders. In D. A. Clark (Ed.), *Unwanted intrusive thoughts in clinical disorders*. New York: Guilford Publications.
- Cramer, P. (1991). *The Development of defense mechanisms: Theory, research, and assessment*. New York, NY: Springer.
- D'Andrade, R. (1987). A folk model of the mind. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 112–148). New York: Cambridge University Press.
- D'Argembeau, A., Renaud, O., & Van der Linden, M. (2011). Frequency, characteristics and functions of future-oriented thoughts in daily life. *Applied Cognitive Psychology, 25*(1), 96–103.
- Davies, M. I., & Clark, D. M. (1998). Thought suppression produces a rebound effect with analogue post-traumatic intrusions. *Behaviour Research and Therapy, 36*(6), 571–582.
- DeWall, C. N., Baumeister, R. F., Gailliot, M. T., & Maner, J. K. (2008). Depletion makes the heart grow less helpful: Helping as a function of self-regulatory energy and genetic relatedness. *Personality and Social Psychology Bulletin, 34*(12), 1653–1662.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69–106.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories: Elaboration and extension of the model. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory, 6*, 322–333.

- Fisher, P. L., & Wells, A. (2005). Experimental modification of beliefs in obsessive-compulsive disorder: A test of the metacognitive model. *Behaviour Research and Therapy*, *43*(6), 821–829.
- Förster, J., & Liberman, N. (2001). The role of attribution of motivation in producing postsuppressional rebound. *Journal of Personality and Social Psychology*, *81*(3), 377.
- Förster, J., & Liberman, N. (2004). A motivational model of post-suppressional rebound. *European review of social psychology*, *15*(1), 1–32.
- Fox, K. C., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R., & Christoff, K. (2015). The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage*, *111*, 611–621.
- Franklin, M. S., Mrazek, M. D., Anderson, C. L., Johnston, C., Smallwood, J., Kingstone, A., et al. (2014). Tracking distraction: The relationship between mind-wandering, meta-awareness, and ADHD symptomatology. *Journal of Attention Disorders*. doi:[10.1177/1087054714543494](https://doi.org/10.1177/1087054714543494)
- Franklin, M. S., Mrazek, M. D., Anderson, C. L., Smallwood, J., Kingstone, A., & Schooler, J. W. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in Psychology*, *4*, e583. doi:[10.3389/fpsyg.2013.00583](https://doi.org/10.3389/fpsyg.2013.00583)
- Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, *11*(4), 303–327.
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*, 325–336.
- Gailliot, M. T., Peruche, B. M., Plant, E. A., & Baumeister, R. F. (2009). Stereotypes and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping. *Journal of Experimental Social Psychology*, *45*(1), 288–290.
- Galera, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribereau-Gayon, R., et al. (2012). Mindwandering and driving: Responsibility case-control study. *British Medical Journal*, *345*, e8105. doi:[10.1136/bmj.e8105](https://doi.org/10.1136/bmj.e8105)
- Giambra, L. M. (1989). Task-unrelated thought frequency as a function of age: A laboratory study. *Psychology of Aging*, *4*, 136–143.
- Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition*, *4*, 1–21.
- Gonzaga, G. C., Haselton, M. G., Smurda, J., Davies, M. S., & Poore, J. C. (2008). Love, desire, and the suppression of thoughts of romantic alternatives. *Evolution and Human Behavior*, *29*(2), 119–126.
- Grodsky, A., & Giambra, L. M. (1990). The consistency across vigilance and reading tasks of individual differences in the occurrence of task-unrelated and task-related images and thoughts. *Imagination, Cognition & Personality*, *10*, 39–52.
- Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, *136*, 495–525.
- Hannan, S. E., & Tolin, D. F. (2005). Mindfulness-and acceptance-based behavior therapy for obsessive-compulsive disorder. In *Acceptance and mindfulness-based approaches to anxiety* (pp. 271–299). Springer US.
- Harrington, J. A., & Blankenship, V. (2002). Ruminative thoughts and their relation to depression and anxiety. *Journal of Applied Social Psychology*, *32*(3), 465–485.
- Harvey, A. G. (2002). A cognitive model of insomnia. *Behaviour Research and Therapy*, *40*(8), 869–893.
- Harvey, A. G., & Payne, S. (2002). The management of unwanted pre-sleep thoughts in insomnia: Distraction with imagery versus general distraction. *Behaviour Research and Therapy*, *40*(3), 267–277.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Holmes, D. (1968). Dimensions of projection. *Psychological Bulletin*, *69*, 248–268.

- Iijima, Y., & Tanno, Y. (2012). The rebound effect in the unsuccessful suppression of worrisome thoughts. *Personality and Individual Differences, 53*(3), 347–350.
- Jacoby, L. L., & Kelley, C. M. (1992). A process-dissociation framework for investigating unconscious influences: Freudian slips, projective tests, subliminal perception, and signal detection theory. *Current Directions in Psychological Science, 1*(6), 174–179.
- Jensen, C. G., Vangkilde, S., Frokjaer, V., & Hasselbalch, S. G. (2012). Mindfulness training affects attention—Or is it attentional effort? *Journal of Experimental Psychology: General, 141*, 106–123.
- Jha, A. P., Morrison, A. B., Dainer-Best, J., Parker, S., Rostrup, N., & Stanley, E. (2015). Minds “at Attention”: Mindfulness training curbs attentional lapses in military cohorts. *PLoS ONE, 10*, e0116889.
- Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion—Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science, 21*, 1686–1693.
- Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2013). Beliefs about willpower determine the impact of glucose on self-control. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 14837–14842.
- Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2015). Beliefs about willpower predict self-regulation and grades in everyday life. *Journal of Personality and Social Psychology, 108*, 637–647.
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice, 10*, 144–156.
- Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18*(7), 614–621.
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932.
- Klinger, E. (2009). Daydreaming and fantasizing: Thought flow and motivation. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds.), *Handbook of imagination and mental simulation* (pp. 225–239). New York, NY: Psychology Press.
- Klinger, E. (2013). Goal commitments and the content of thoughts and dreams: Basic principles. *Frontiers in Psychology, 4*, e415. doi:10.3389/fpsyg.2013.00415
- Klinger, E., & Cox, W. M. (1987). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality, 7*(2), 105–128.
- Koster, E. H., Rassin, E., Crombez, G., & Näring, G. W. (2003). The paradoxical effects of suppressing anxious thoughts during imminent threat. *Behaviour Research and Therapy, 41*(9), 1113–1120.
- Lambert, A. E., Hu, Y., Magee, J. C., Beadel, J. R., & Teachman, B. A. (2014). Thought suppression across time: Change in frequency and duration of thought recurrence. *Journal of Obsessive-Compulsive and related Disorders, 3*(1), 21–28.
- Levinson, D. B., Smallwood, J., & Davidson, R. J. (2012). The persistence of thought: Evidence for a role of working memory in the maintenance of task-unrelated thinking. *Psychological Science, 23*, 357–380.
- Logel, C., Iserman, E. C., Davies, P. G., Quinn, D. M., & Spencer, S. J. (2009). The perils of double consciousness: The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology, 45*(2), 299–312.
- Magee, J. C., Harden, K. P., & Teachman, B. A. (2012). Psychopathology and thought suppression: A quantitative review. *Clinical Psychology Review, 32*(3), 189–201.
- Magee, J. C., & Teachman, B. A. (2007). Why did the white bear return? Obsessive-compulsive symptoms and attributions for unsuccessful thought suppression. *Behaviour Research and Therapy, 45*(12), 2884–2898.
- Malle, B., & Knobe, J. (1997). Which behaviors do people explain? A basic actor-observer asymmetry. *Journal of Personality and Social Psychology, 72*, 288–304.
- Marcks, B. A., & Woods, D. W. (2005). A comparison of thought suppression to an acceptance-based technique in the management of personal intrusive thoughts: A controlled evaluation. *Behaviour Research and Therapy, 43*(4), 433–445.

- Marcks, B. A., & Woods, D. W. (2007). Role of thought-related beliefs and coping strategies in the escalation of intrusive thoughts: An analog to obsessive-compulsive disorder. *Behaviour Research and Therapy*, *45*(11), 2640–2651.
- Martijn, C., Tenbült, P., Merckelbach, H., Dreezens, E., & de Vries, N. K. (2002). Getting a grip on ourselves: Challenging expectations about loss of energy after self-control. *Social Cognition*, *20*, 441–460.
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*, 196–204.
- McVay, J. C., & Kane, M. J. (2010). Does mind-wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*, 188–197.
- McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, *16*, 857–863.
- Molden, D. C., & Dweck, C. S. (2006). Finding “meaning” in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, *6*, 192–203.
- Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*, 11–18. doi:10.1037/a0031569
- Morewedge, C. K., Giblin, C. E., & Norton, M. I. (2014). The (perceived) meaning of spontaneous thoughts. *Journal of Experimental Psychology: General*, *143*(4), 1742.
- Morsella, E., Ben-Zeev, A., Lanska, M., & Bargh, J. A. (2010). The spontaneous thoughts of the night: How future tasks breed intrusive cognitions. *Social Cognition*, *28*(5), 641.
- Mrazek, A. J., Ihm, E., Molden, D. M., Mrazek, M. D., Zedelius, C., & Schooler, J. W. (2017). *Growth mindsets of self-control and the role of effort*. Preliminary data, University of California, Santa Barbara
- Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, *24*, 774–781.
- Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: Finding convergence through opposing constructs. *Emotion*, *12*, 442–448.
- Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*, 147–259.
- Murray, H. A. (1951). Uses of the thematic apperception test. *American Journal of Psychiatry*, *107* (8), 577–581.
- Najmi, S., Riemann, B. C., & Wegner, D. M. (2009). Managing unwanted intrusive thoughts in obsessive-compulsive disorder: Relative effectiveness of suppression, focused distraction, and acceptance. *Behaviour Research and Therapy*, *47*, 494–503.
- Obsessive Compulsive Cognitions Working Group. (2005). Psychometric validation of the obsessive belief questionnaire and interpretation of intrusions inventory—Part 2: Factor analyses and testing of a brief version. *Behaviour Research and Therapy*, *43*(11), 1527–1542.
- Owens, G., Chard, K., & Cox, T. (2008). The relationship between maladaptive cognitions, anger expression, and posttraumatic stress disorder among veterans in residential treatment. *Journal of Aggression, Maltreatment & Trauma*, *17*, 439–452. doi:10.1080/10926770802473908
- Plaks, J. E., Grant, H., & Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception. *Journal of Personality and Social Psychology*, *88*(2), 245.
- Plaks, J. E., Levy, S. R., & Dweck, C. S. (2009). Lay theories of Personality: Cornerstones of meaning in social cognition. *Social and Personality Compass*, *3*, 1069–1081.
- Poole, D. A., Lindsay, D. S., Memon, A., & Bull, R. (1995). Psychotherapy and the recovery of memories of childhood sexual abuse: US and British practitioners’ opinions, practices, and experiences. *Journal of Consulting and Clinical Psychology*, *63*(3), 426.

- Posner, M. I., Rothbart, M. K., & Tang, Y.-Y. (2015). Enhancing attention through training. *Current Opinion in Behavioral Sciences*, 4, 1–5.
- Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, 111(3), 781.
- Purdon, C., & Clark, D. A. (2002). The need to control thoughts. In R. O. Frost & G. Steketee (Eds.), *Cognitive approaches to obsessions and compulsions: Theory, assessment, and treatment*. London: Elsevier Science.
- Rachman, S. (1981). Part I. Unwanted intrusive cognitions. *Advances in Behaviour Research and Therapy*, 3(3), 89–99.
- Rachman, S. (1997). A cognitive theory of obsessions. *Behaviour Research and Therapy*, 35(9), 793–802.
- Rachman, S., & de Silva, P. (1978). Abnormal and normal obsessions. *Behaviour Research and Therapy*, 16(4), 233–248.
- Radomsky, A. S., Alcolado, G. M., Abramowitz, J. S., Alonso, P., Belloch, A., Bouvard, M., et al. (2014). Part 1—You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders*, 3(3), 269–279.
- Reuven-Magril, O., Rosenman, M., Liberman, N., & Dar, R. (2009). Manipulating meta-cognitive beliefs about the difficulty to suppress scratching: Implications for obsessive-compulsive disorder. *International Journal of Cognitive Therapy*, 2(2), 143–151.
- Reynolds, M., & Brewin, C. R. (1999). Intrusive memories in depression and posttraumatic stress disorder. *Behaviour Research and Therapy*, 37(3), 201–215.
- Reynolds, M., & Salkovskis, P. M. (1992). Comparison of positive and negative intrusive thoughts and experimental investigation of the differential effects of mood. *Behaviour Research and Therapy*, 30(3), 273–281.
- Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology*, 26, 234–242.
- Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind wandering to situational demands. *Psychonomic Bulletin & Review*, 21, 1309–1315.
- Sarason, I. G., Pierce, G. R., & Sarason, B. R. (1996). Domains of cognitive interference. In I. G. Sarason, G. R. Pierce, & B. R. Sarason (Eds.), *Cognitive interference: Theories, methods, and findings* (pp. 139–152). Mahwah, NJ: Erlbaum.
- Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, 1(1), 60.
- Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences*, 6, 339–344.
- Schooler, J. W., Mrazek, M. D., Franklin, M. S., Baird, B., Mooneyham, B. W., Zedelius, C., et al. (2014). The middle way: Finding the balance between mindfulness and mind-wandering. *Psychology of learning and motivation*, 60, 1–33.
- Schooler, J. W., Reichele, E., & Halpern, D. (2004). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children*, Cambridge, MA.
- Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Neurosciences*, 15, 319–326. doi:10.1016/j.tics.2011.05.006
- Scruton, R. (1982). *Kant*. Oxford: Oxford University Press.
- Sedlmeier, P., Eberth, J., Schwarz, M., Zimmermann, D., Haerig, F., Jaeger, S., et al. (2012). The psychological effects of meditation: A meta-analysis. *Psychological Bulletin*, 138, 1139–1171.
- Seli, P., Carriere, J. S. A., & Smilek, D. (2014). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. *Psychological Research*, e617. doi:10.1007/s00436-014-0617-x
- Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*.

- Shafran, R., & Rachman, S. (2004). Thought-action fusion: A review. *Journal of Behavior Therapy and Experimental Psychiatry*, 35(2), 87–107.
- Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought-action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders*, 10, 379–391.
- Singer, J. L., & Schonbar, R. A. (1961). Correlates of daydreaming: A dimension of self-awareness. *Journal of Consulting Psychology*, 25, 1.
- Smallwood, J. (2010). Why the global availability of mind-wandering necessitates resource competition: Reply to McVay and Kane (2010). *Psychological Bulletin*, 136, 202–207.
- Smallwood, J., McSpadden, M. C., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, 14, 527–533.
- Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory and Cognition*, 36, 1144–1150.
- Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition*, 18(1), 118–125.
- Smallwood, J., Obonsawin, M., & Reid, H. (2003). The effects of block duration and task demands on the experience of task unrelated thought. *Imagination, Cognition and Personality*, 22, 13–31.
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, 132, 946–958.
- Stern, R. (2004). Does 'ought' imply 'can'? And did Kant think it does? *Utilitas*, 16(01), 42–61.
- Tang, Y.-Y., & Posner, M. I. (2009). Attention training and attention state training. *Trends in Cognitive Sciences*, 13, 222–227.
- Vranas, P. B. (2007). I ought, therefore I can. *Philosophical Studies*, 136(2), 167–216.
- Wammes, J. D., Seli, P., Allan, J., Boucher, P. O., & Smilek, D. (2016). Mind wandering during lectures II: Relation to academic performance. *Scholarship of Teaching and Learning in Psychology*, 2(1), 33–48.
- Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, 101, 34–52.
- Wegner, D. M. (2011). Setting free the bears: Escape from thought suppression. *American Psychologist*, 66(8), 671.
- Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, 63(6), 903.
- Wegner, D. M., & Gold, D. (1995). Fanning old flames: Emotional and cognitive effects of suppressing thoughts of a past relationship. *Journal of Personality and Social Psychology*, 68, 782–792.
- Wegner, D. M., & Pennebaker, J. W. (1993). Changing our minds: An introduction to mental control. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of mental control* (pp. 1–12). Englewood Cliffs, NJ: Prentice-Hall.
- Wegner, D. M., & Smart, L. (1997). Deep cognitive activation: A new approach to the unconscious. *Journal of Consulting and Clinical Psychology*, 65(6), 984–995.
- Wells, A. (2005). Detached mindfulness in cognitive therapy: A metacognitive analysis and ten techniques. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, 23(4), 337–355.
- Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology*, 51(1), 59–91.
- Wenzlaff, R. M., Wegner, D. M., & Roper, D. W. (1988). Depression and mental control: The resurgence of unwanted negative thoughts. *Journal of Personality and Social Psychology*, 55(6), 882.
- Wicklow, A., & Espie, C. A. (2000). Intrusive thoughts and their relationship to actigraphic measurement of sleep: Towards a cognitive model of insomnia. *Behaviour Research and Therapy*, 38(7), 679–693.
- Wilkinson-Tough, M., Bocci, L., Thorne, K., & Herlihy, J. (2010). Is mindfulness-based therapy an effective intervention for obsessive-intrusive thoughts: A case series. *Clinical Psychology & Psychotherapy*, 17(3), 250–268.

- Williams, E. F., & Gilovich, T. (2008). Do people really believe they are above average? *Journal of Experimental Social Psychology, 44*(4), 1121–1128.
- Wolpe, J. (1990). *The practice of behavioral therapy* (4th ed.). New York, NY: Pergamon Press.
- Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 301–314.
- Yeager, D. S., & Walton, G. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267–301.
- Zedelius, C. M., Broadway, J. M., & Schooler, J. W. (2015). Motivating meta-awareness of mind wandering: A way to catch the mind in flight? *Consciousness and Cognition, 36*, 44–53.
- Zedelius, C. M., Protzko, J., & Schooler, J. W. (2017a). *Below-average lost in thoughts: People believe their own mind wandering rates to be lower than those of other people*. Preliminary data, University of California, Santa Barbara.
- Zedelius, C. M., Protzko, J., & Schooler, J. W. (2017b). *Lay theories of mind wandering affect the rate of mind wandering in everyday life and in the lab*. Preliminary data, University of California, Santa Barbara.
- Zedelius, C. M., & Schooler, J. W. (2015). Mind wandering “Ahas” versus mindful reasoning: Alternative routes to creative solutions. *Frontiers in Psychology, 6*.
- Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: Evidence of brief mental training. *Consciousness and Cognition, 19*, 597–605.

Lay Theories of Creativity

Simone M. Ritter and Eric F. Rietzschel

Creativity is considered to be one of humanity's most complex and important behaviors, and its effects are widespread. Over time, creativity has allowed us to create art, develop computers, and cure illnesses. In addition to its importance in science and the arts (Feist & Gorman, 1998; Kaufman, 2002; Mackinnon, 1962), the significance of creativity has also been recognized in daily life problem-solving (Cropley, 1990) and in successful adaptation to change (Cropley, 1990; Reiter-Palmon, Mumford, & Threlfall, 1998). Moreover, creativity helps us to sustain and promote our well being (Hirt, Devers, & McCrea, 2008), it allows us to gain power (Sligte, De Dreu, & Nijstad, 2011), it makes us more attractive mating partners (Griskevicius, Cialdini, & Kenrick, 2006), and it is core to successful innovation (Amabile, 1996). Due to the crucial role of creativity in innovation, creativity has become a key concern for most organizations and businesses (Runco, 2004), and some scholars even refer to today's economy as a *creative economy* (Florida, 2002; Howkins, 2002).

Not only in our current society, but throughout the history of mankind, creativity has been of great appeal and importance to people (Mithen, 1998). Therefore, it is not surprising that people strive to understand creativity, and that they develop *lay theories* to do so. Lay theories are the informal theories and beliefs that lay people hold about a phenomenon and its causes or consequences (e.g., Furnham, 1988). Lay theories are usually not based on scientific research or a systematic analysis of the phenomenon in question (although some lay theories may be partly informed by such research), but rather take the shape of stereotypes and everyday beliefs

S. M. Ritter (✉)

Behavioural Science Institute, Radboud University Nijmegen,
P.O. Box 9104, 6500 HE Nijmegen, The Netherlands
e-mail: S.Ritter@psych.ru.nl

E.F. Rietzschel

Department of Psychology, University of Groningen, Groningen, The Netherlands
e-mail: e.f.rietzschel@rug.nl

(‘everybody knows that ...’). Lay theories can be encountered in everyday conversations and popular publications (books, news articles, websites). However, they are not always formulated as an explicit theory, but sometimes take the form of unspoken assumptions, or implicit beliefs. As we will describe in this chapter, lay theories about creativity have been developed about personality traits and mental disorders that are associated with creativity, about what can be considered to be creative, and about the specific environments or techniques that support or kill creativity.

The lay theories and beliefs that people hold about creativity (i.e., everyday, informal, and often implicit beliefs about creativity, how it works, how it is best stimulated, or who has the highest creative potential) are not just important from a theoretical perspective, but may directly influence creativity itself (Runco, Johnson, & Baer, 1993). For example, beliefs about personality traits and characteristics associated with creativity influence whether people, ideas, and products are perceived and recognized as creative, and beliefs about the creative processes shape what parents, educators, and organizations do in order to facilitate creativity.

Research conducted by Baas, Koch, Nijstad, and De Dreu (2015) demonstrated that people have strong beliefs about creativity and, importantly, that these beliefs are often incomplete and not in line with the state-of-the-art scientific evidence. This becomes problematic when such lay beliefs inform the choices that people make. Lay beliefs, for example, shape the circumstances people create in order to stimulate their and each other’s creativity (Baas et al., 2015). They may, thus, lead stakeholders such as policy makers, supervisors, and instructors to develop or implement expensive but ineffective—and in the worst case even harmful—interventions. Therefore, a better understanding on which beliefs about creativity are supported by scientific evidence, and which can be considered misunderstandings, will help to foster creativity in the entire population. The aim of the current chapter is to test several lay theories about creativity against the available scientific evidence. We will describe these lay theories and beliefs, and then critically appraise them in light of what creativity research has shown.

Lay Theories of Creativity

Unfortunately, there is little systematic research on the content or the structure of the lay theories that people hold about creativity. However, there has been some research about specific beliefs and stereotypes that people hold, and we will discuss these in the current chapter—as we will see, creativity researchers have worked hard to dispel some of these beliefs and stereotypes. Moreover, we will address various lay theories and lay beliefs about creativity that are clearly present in the field, that is, among people who work on a professional or semi-professional basis in the field of creativity, or in creative professions. One can encounter these lay theories, for example, when reading blogs or books about creativity, when visiting websites of companies that offer creativity trainings, business seminars, or creativity

and innovation consultancy, or when browsing through creativity-relevant quotes by famous creative individuals. However, few of these can really be counted as ‘theories,’ in the sense that they are used systematically to explain or predict. Furnham (1988) describes several characteristics of lay theories as compared to scientific theories, such as their lack of explicitness (people may not always be able to clearly state their theories), lack of coherence and consistency (the theories may be fragmented and self-contradictory), the emphasis on verification rather than falsification (people are generally more interested in applying their lay theories than in testing them), and the focus on content rather than process (i.e., lay theories tend to describe types or categories, rather than describe underlying processes that may give rise to certain differences). It seems that this also holds for lay theories or lay beliefs regarding creativity. Generally speaking, then, we will use the term ‘lay theories’ rather loosely, as referring to all theories, beliefs, or stereotypes that lay people tend to hold regarding creativity.

One way of organizing these lay theories is by using the ‘Four P’ model of creativity, which distinguishes between the creative *Person*, *Process*, *Press (or Place)*, and *Product* (Rhodes, 1961). This principle is commonly used to organize the research literature on creativity; we will use it to organize the different lay theories we discuss (and the support or lack thereof). Thus, lay theories concerning the following aspects of creativity will be addressed: what are individual characteristics of the creator (*Person*), which skills and processes are needed to achieve creativity (*Process*), which environment stimulates or hinders creativity (*Place*), and what is considered to be creative (*Product*).

Defining Creativity

Before we compare lay theories about creativity with scientific knowledge about the creative person, process, press, and product, we should provide a *definition of creativity*. Creativity can take many forms and can be found within a variety of contexts. What exactly is creativity? The word *creativity* has its roots in the Latin term *creō*, which means ‘to create, to make.’ In the current chapter, we use a definition of creativity that is commonly used in the research literature: creativity is the ability to generate ideas or problem solutions that are both *original* and *useful* (e.g., Amabile, 1983; Mumford, 2003; Sternberg & Lubart, 1999). If something is not novel, unusual, or unique, it is mundane, commonplace, or conventional—it is not original, and therefore not creative. Important to notice, an idea or product that is original can be unique or uncommon for a good reason: it might be useless. Ideas or responses that are highly unusual, but not appropriate to the task at hand, might be called eccentric, bizarre, or even pathological, but not creative. In other words, originality is vital (Barron, 1955; Stein, 1953), but must be balanced with ‘usefulness,’ that is, with fit and appropriateness (Runco, 1988).

Lay Theories About the Creative Person

Some of the best-known lay theories on creativity revolve around the creative person: who are those creative people, and what makes them so special? One global implicit theory underlying these questions is that there is not only a clear distinction between creative and less creative people (i.e., that the former have distinct abilities or traits that they do not share with the rest of us), but also that creative potential is something one either has or does not have, without much room for improvement. In this section, we will discuss three (interrelated) lay theories on creativity: the ‘creative genius’ belief, the ‘mad genius’ belief, and the belief that older people are less creative.

Creative Genius

The natural starting point for enquiries into creativity is to think about *examples* of creativity: when we think about creative performance, what comes to mind? Following this availability heuristic, the exemplars that often come to mind are well-known cases of extremely gifted creators, who have had a major and lasting impact on their field (and sometimes even outside of their field); names like Mozart, Da Vinci, or Einstein are among the ones people tend to mention. The association between the concept of ‘creativity’ and these eminent exemplars seems to be quite strong, and indeed, in the absence of modern research methods and detailed theories of the creative process, early creativity research focused on case studies of eminent creative individuals and their work (e.g., Ghiselin, 1952; Guilford, 1950; see Van Strien, 2015, for a recent overview and discussion). Although creativity research has since then expanded its focus considerably, research on eminent creative individuals is still important (e.g., Simonton, 2004).

The ‘creative genius’ theory comprises three different beliefs, which we will discuss in turn: first, the belief that *creativity is rare*; second, the belief that this kind of *creativity is fundamentally different* from everyday creativity; and third, an ‘entity theory’ regarding creative potential, stating that *creativity is unlearnable*.

Is Creativity Rare?

The first issue basically comes down to a point of semantics, and whether one wishes to limit the meaning of the term ‘creativity’ to exceptional cases. In fact, people often tend to adopt this narrow meaning, either implicitly or explicitly; in (informal) discussions of creativity, it usually does not take long before somebody brings up Einstein or Mozart (or perhaps a more contemporary example like Steve Jobs), and makes the argument that such exceptional cases reflect ‘true’ creativity. Indeed, in line with Furnham’s (1988) discussion, lay theories of creativity often

seem to be strongly categorical, in that people tend to want to arrive at some criterion to decide whether somebody (or something, in case of a creative product) ‘is creative or not.’ In the creativity literature, in contrast, researchers generally distinguish between different *types* or *levels* of creativity. For example, a distinction is often made between so-called *Big-C* and *little-c* creativity (e.g., Csikszentmihalyi, 1997; Gardner, 1993). *Big-C* creativity refers to highly eminent creators who have had a lasting and transformative impact on their field; *little-c* creativity refers to the rest of us, people who may display creativity in their own way, but who will never reach the level of the truly great. Kaufman and Beghetto (2009) extend this distinction into a Four-C model, adding *mini-c* creativity (individual creative insights) and *pro-C* creativity (creative performance at a highly skilled, professional level that is nevertheless not revolutionary or transformative). Similarly, Boden (2004) distinguishes between *P-creativity* and *H-creativity*. *P-creativity* is *psychological* creativity, creative ideas or insights that are novel to the person who had the idea; *H-creativity* is *historical* creativity, referring to ideas that are novel for humanity or society in general. Thus, one of the main differences between lay and scientific theories of creativity is that the latter do not consider exceptional and lasting impact as a criterion to decide whether somebody is creative or not; instead, scientific creativity theories acknowledge that creativity can occur at different levels.

Are Creative People Fundamentally Different?

Whichever distinction one uses, the question remains whether there is anything *fundamentally* different between these levels of creative performance; i.e., whether what *Big-C* creators do, differs in a qualitative sense from what all others do, or whether the processes or abilities leading to *H-creativity* are fundamentally different from those leading to *P-creativity*. Again, this is a very popular notion. A Google search for ‘highly creative people’ yields many articles with titles such as “20 Things Only Highly Creative People Would Understand” (Kaiser, n.d.), “18 Things Highly Creative People Do Differently” (Gregoire, 2016), and “Creative People’s Brains Really Do Work Differently” (Gregoire & Kaufman, 2016). What such articles have in common is that, even though many of them are grounded in actual research, they strongly emphasize the difference between highly creative people and the rest of humanity, suggesting that there really is some fundamental difference between ‘the creative person’ and ordinary people. Van Strien (2012) notes that this belief can be traced back to ancient Greece, with its conception of the artist as somebody who is susceptible to a form of divine inspiration not available to ordinary people, and further shows how the belief of the creative person as ‘extraordinary’ (in a literal sense) was further cultivated in the Romantic period. There seems to be a persistent belief that creative individuals possess a set of traits or abilities that somehow sets them apart from the rest of humanity. However, the consensus in the research field is that this is not the case, and that even creative performance at exceptional levels differs gradually, not fundamentally, from other

levels of creativity (the *continuity principle*; Guilford, 1950). Although Big-C creativity may require a unique *combination* of individual and contextual factors that only rarely co-occur at high levels (e.g., high abilities, strong internal motivation, relevant personality traits, opportunities for training, and interpersonal skills), none of these traits in themselves are fundamentally different from those of other people who perform at more ordinary levels. Thus, for example, in the next section (Creative Process), we will address the *creative cognition* approach, which explicitly takes the continuity principle as its starting point (e.g., Finke, Ward, & Smith, 1995; Nijstad & Stroebe, 2006; Ward, 1994).

Is Creativity Unlearnable?

A third aspect of the creative genius theory is that, since the difference between highly creative and less creative people is fundamental, creativity is not something we can learn; we either have it, or we do not. Although few if any informal sources espouse this belief (in fact, most intend to help people *overcome* this idea), it is a common remark in informal conversations about creativity: many people seem to believe that they “simply are not creative.” These negative claims are then usually supported by ‘evidence’ such as a lack of artistic skills. This kind of thinking is what Dweck, Hong, and Chiu (1993) refer to as an *entity theory*: the belief that performance is a matter of stable, unchangeable traits. Entity theories are contrasted with *incremental theories*, according to which performance is something that can be enhanced through development of the underlying traits and abilities. Thus, if someone fails to perform creatively, an entity theorist might conclude that this person simply ‘is not creative,’ whereas an incrementalist might conclude that he or she needs to develop certain creativity-relevant skills (e.g., Amabile, 1996). As we will see later, the risk of entity theories is that they can become a self-fulfilling prophecy, because they tend to significantly inhibit motivation and learning performance as compared to incremental theories (Burnette, O’Boyle, VanEpps, Pollack, & Finkel, 2013).

Of course, even if creativity is considered a stable trait, this does not mean that it is impossible or unnecessary to try to stimulate it. Even stable traits, such as personality dimensions, are associated with a broad range of intrapersonal variability (Fleeson, 2001): somebody who is highly extraverted, for example, may be led to behave in a more introverted manner in various situations. Thus, the question may not be whether creative potential is stable, and how high one’s level of creative potential is, but rather how *broad* one’s range of intra-individual variation is (Fleeson, 2001). The degree to which creative potential is expressed does, to a certain extent, depend on the situation—extensive research has shown that people can be induced to behave or perform more creatively by a variety of task manipulations and environmental factors. For example, people’s creativity can be stimulated by traveling and other ‘diversifying experiences’ (e.g., Maddux & Galinsky, 2009; Ritter et al., 2012a), by situationally inducing a so-called *promotion focus* (a focus on attaining gains and realizing ambitions; Friedman & Förster, 2001;

Higgins, 1997), and by exposure to creative examples or models (Shalley & Perry-Smith, 2001). Thus, even if some people do have more creative potential than others, there is sufficient evidence for intra-individual variability, and the question can be raised whether it is possible to develop one's creative skills. In fact, researchers agree that creativity *is* something that can be developed, for example, through training (e.g., Ritter & Mostert, 2016; Scott, Leritz, & Mumford, 2004). Moreover, study of typical Big-C creators shows that these people actually spent enormous amounts of time and effort into developing their creative and domain-relevant skills (cf. Ghiselin, 1952), showing that the entity theory of creative ability does not even hold for those examples that are most often invoked in support of the creative genius myth.

Creativity and Psychopathology

Perhaps the most pervasive belief about the creative person, next to the 'creative genius' belief, is the idea that highly creative individuals tend to be unstable at best, and tend to have psychotic tendencies at worst (e.g., Baas, Nijstad, Boot, & De Dreu, 2016; Silvia & Kaufman, 2010; Simonton, 2014a, b; Van Strien, 2015). This belief, traced back as far as ancient Greece and Rome, partly rests on the creative genius belief, in that it seems to be largely based on the availability of highly salient examples of eminent yet unstable creators. Vincent Van Gogh probably is one of the best-known examples of the 'unstable artist,' and as remarked above, the availability of such examples has probably contributed greatly to the stereotype. It may also be partly due to the romantic notion of the artist as a highly sensitive individual, who is in touch with his or her inner life and emotional turmoil to a far greater degree than ordinary people (Van Strien, 2012). Yet another reason for this belief may be that creativity is strongly associated with spontaneity, impulsivity, and a rejection of social or group norms (e.g., Feist, 1998), all of which can, when present at extreme levels, be associated with psychopathology as well.

The belief that highly creative people are mentally unstable has been the subject of debate in the scientific literature; Simonton (2014a) called it the 'mad genius controversy,' with some researchers supporting this belief, and others strongly rejecting it (see Baas et al., 2016, for an overview). To the extent that research data support either the one or the other perspective, the lack of true experiments in this area makes the data difficult to interpret: the available data are typically correlational, precluding causal interpretations. As Furnham (1988) noted, lay theories tend to confuse correlation and causation, and the mad genius belief seems to be a good example of this tendency. Even if there is a correlation between psychopathology and creativity, this does not mean that people are more creative *because* of their psychopathology. First, the causal relation might run in the opposite direction—creativity might somehow contribute to psychopathology, for example, because people get used to taking highly uncommon perspectives and get estranged from 'typical' trains of thought. Second, the relation might be explained

by a third variable that predicts both creativity and a propensity toward psychopathology. Third, even if there is a causal path from psychopathology to creativity, this might exist for other reasons than commonly thought; for example, people might pursue creative endeavors as a coping strategy (e.g., Greene, 1980).

The Role of Approach and Avoidance

Baas et al. (2016) recently published a review and meta-analysis on the mad genius belief. Their paper is worth describing in some detail, as it is one of the few studies that systematically address a specific lay belief in the area of creativity. Baas et al. argue that the relation between creativity and psychopathology is best understood from the perspective of *approach* versus *avoidance motivation* (Carver, Sutton & Scheier, 2000), because—they argue—these reflect two fundamental motivational systems that have been linked to both creativity and various forms of psychopathology. The *approach system* is concerned with eager striving toward (approaching) positive, rewarding outcomes and situations, including novel stimuli and experiences. In contrast, the *avoidance system* is concerned with vigilance, fear, and withdrawal from aversive outcomes and risky situations. Approach motivation has been shown to be a positive predictor of cognitive flexibility and creativity; avoidance motivation, in contrast, generally negatively predicts creativity (however, see Roskes, De Dreu, & Nijstad, 2012, for a possible exception). Baas and colleagues further state that several psychopathologies have strong roots in these approach and avoidance systems; for example, *depression* is linked to avoidance motivation and its correlates, whereas *bipolar disorder* (specifically, mania or hypomania, which are states of high cognitive activation and extremely elated mood states) tends to be associated more with the approach system. In their review and meta-analysis, Baas and colleagues indeed find that approach-related pathologies, such as mania and hypomania, positively relate to creativity, whereas avoidance-related pathologies, such as depression, negatively relate to creativity. However, they also find that the effect sizes for the latter (avoidance-related) relationships are quite small, bordering on the ‘trivial.’ Thus, the relationships are there, but explain only very little variance in creativity.

Creativity and Age

Another pervasive belief is that creativity mostly comes from younger people (Rietzschel, Zacher, & Stroebe, 2016). Actually, this lay belief comprises two different beliefs: first, the belief that children are more creative than adults (at least up to a certain age), and second the belief that adults become less creative in old age.

Children and Creativity

A popular lay belief holds that young children are much more creative than adults, and even than older children. Creativity is thought to decline when children are socialized into thinking along more conventional lines and to worry more about being evaluated positively by others. For example, the creativity website *Creating Minds* states that, “Our decline in creativity does not start when we are 40 or 50. It starts around about the age when we enter school” (*Creating Minds*, n.d.) and goes on to state that “At around about the age of five, we are using about 80% of our creative potential ... by the age of twelve, our creative output has declined to about 2% of our potential, and it generally stays there for the rest of our lives.” A related claim is made by the website *stephenshapiro*, which states that “98% of 5-year olds test as highly creative, yet only 2% of adults do.” Other such claims are not difficult to find, as a Google search for ‘children more creative than adults’ will show. What is difficult, however, is to find scientific research actually supporting them. It is not always clear where the numbers come from, or even what they are supposed to mean, for example, if 98% of children “test as highly creative,” this should raise some serious concerns regarding the norm scores used.

Regardless of such issues, however, the basis for the “children are more creative” belief seems to lie mostly in the notion that young children supposedly do not worry as much about giving the ‘correct’ response, and instead are more likely to do what they like or what occurs to them. As they get older, children learn that some behaviors are rewarded, for example, with praise or with high grades, and this—it is thought—guides their development away from creativity. In a way, this reasoning actually is in line with some of what we know from research. For example, research has shown that creativity can easily be inhibited by evaluation and rewards (we will return to this below, in the section on the creative environment), and that even young children who ‘learn’ to do a creative task (such as making a drawing) in order to gain a reward (such as a piece of candy or the opportunity to play with an attractive toy) often perform less creatively, and enjoy the task less (see Amabile, 1996, for an overview). However, whether this means that children are more creative than adults in a direct comparison, or whether such a direct comparison even makes sense to begin with, is far from clear.

Older Adults and Creativity

In general, older people are confronted with a variety of negative stereotypes (Cuddy, Norton, & Fiske, 2005; Fiske, Cuddy, Glick, & Xu, 2002; Lamont, Swift, & Abrams, 2015), and creativity is no exception. Young age tends to be associated with curiosity, flexibility, energy, and creativity, whereas old age tends to be associated with a lack of flexibility, lower motivation, decreased cognitive abilities, and a lack of openness to change (Ng & Feldman, 2012; Shearring, 1992). Whether this is a specific lay belief about age and creativity or simply an extension of the general belief that old age comes with a general decline in physical and

psychological abilities is difficult to say, but it is clear that older people are generally thought to be less creative than their younger counterparts.

Empirical research in fact does not show a clear disadvantage of old age for creativity. In a recent review article on the relation between age and creativity at work, Rietzschel et al. (2016) reviewed two meta-analyses and seven further primary studies on the relation between age and creativity and innovation at work. They conclude that most research shows no direct relation between age and creativity. For example, a meta-analysis by Ng and Feldman (2008) found no significant relationships between age and either self-reported or supervisor-rated creativity and innovation. Further, a more recent meta-analysis by Ng and Feldman (2013) on the relation between age, job tenure, and innovation-related behaviors showed no relations between age and innovation-related behaviors, with the exception of a weak positive relation between age and self-rated innovative behavior. As a concrete example, Rietzschel et al. (2016) also give a short summary of research on age and scientific creativity. Again, research suggests that older people (in this case, scientists) are no less creative than their younger colleagues; although there is evidence for a curvilinear relation between age and scientific productivity (peaking around the age of 40–45; Stroebe, 2010; also see Simonton, 1997), this seems to have changed in the last two decades. For example, a study by Gingras, Lariviere, Macaluso, and Robitaille (2008) failed to find a significant decrease in productivity even after the age of 50 years.

Lay Theories About the Creative Process

As explained earlier, lay beliefs often assume that creativity is a matter of innate talent that only a few people possess. Most creativity researchers, however, agree that highly eminent creativity (Big-C creativity) is *not* fundamentally different from ordinary creativity (little-c creativity), that is, they rely on a continuity principle (Guilford, 1950). For example, the creative cognition approach defines creativity as the product of ‘ordinary’ cognitive processes that are used to produce something extraordinary (e.g., Smith, Ward, & Finke, 1995; Nijstad & Stroebe, 2006; Ward, 1994). Using models and concepts from ‘ordinary’ cognitive psychology, studies in the creative cognition tradition have been able to study, predict, and explain creative performance quite well (e.g., Smith et al., 1995; Nijstad & Stroebe, 2006; Ward, 1994). Basal examples for the inborn talent to create are the flexible use of language—through which we can generate a tremendous variety of novel constructions (Chomsky, 1972; Pinker, 1984)—the ability to combine concepts to generate more complex ones, and the capacity to map properties analogically across different domains (Finke, Ward, & Smith, 1992). These processes are, in themselves, creative and, moreover, they underlie all forms of creativity—from the most mundane to the most extraordinary. Important to notice, although creative thinking skills are considered normal cognitive functions, individual differences in creativity exist, for example, due to variations in the use of specific processes, deviation in the

intensity of process use, and differences in the combination of processes. Moreover, many researchers agree that creative skills can, to some extent, be nurtured (Ritter, Strick, Bos, van Baaren, & Dijksterhuis, 2012c; Scott et al., 2004).

Divergent and Convergent Creativity

A common lay belief is that creative thought equals *divergent* thought. This seems to stem from the belief that creativity is fundamentally different from ‘ordinary’ behavior and cognition, and that those kinds of cognitions and behaviors that get formally taught and rewarded in society must, by extension, be incompatible with creativity. Although the creative process involves divergent thought, they are not synonymous. In fact, the creative process entails both divergent thought and convergent thought (Guilford, 1967; Maier, 1967; Simon, Newell, & Shaw, 1962). *Divergent thought* involves producing multiple or alternative answers from available information by making unexpected combinations, recognizing links among remote associates, or transforming information into unexpected forms. A typical example of divergent creativity is idea generation (e.g., during a brainstorming session). There is strong evidence which suggests that divergent thinking represents a distinct ability necessary for many forms of creative performance (Bachelor & Michael, 1991, 1997; Mumford, Marks, Connelly, Zaccaro, & Johnson, 1998; Plucker & Renzulli, 1999; Scott et al., 2004; Scratchley & Hakstian, 2001; Sternberg & O’Hara, 1999; Vincent, Decker, & Mumford, 2002). However, although important, divergent thought is only one component of the creative process. Many scholars emphasize the need for an additional cognitive ability, *convergent thinking*. Convergent thought is the cognitive process of deriving the single best, or most correct, answer to a problem or question (Fasko, 2001; Nickerson, 1999; Treffinger, 1995). Convergent thought emphasizes accuracy and logic, and applies conventional search, recognition, and decision-making strategies, and as such can easily be considered to be ‘uncreative,’ but it may actually still require creativity as well. Convergent thought, for example, is required in tasks where seemingly unrelated concepts have to be related, as measured in the *Remote Associates Test* (Mednick, 1962). In this task one has to generate a fourth word, which connects three seemingly unrelated words (example: bar—dress—glass, fourth word: cocktail; cocktail bar, cocktail dress, cocktail glass). Further, convergent thought and divergent thought are often combined. A creative activity that requires the strong interplay of divergent and convergent thought is *creative problem-solving*—the cognitive process of searching for a novel and inconspicuous solution to a problem. For example, in the two-string problem, participants are required to tie together two strings hanging from the ceiling. However, the strings are arranged so far apart that they cannot be reached at the same time. The solution requires the use of one of the objects available in the room so that one string can be set in motion as a pendulum. This swinging string can then be caught, while holding the other string, and thus they can be tied together. To solve this problem,

divergent thought is needed to come up with the idea to use the displayed object in an unfamiliar manner, and convergent thought is needed to verify the problem solution.

Creativity and Flexibility

In a recent study, Baas et al. (2015) asked laypeople which processes they believed to be beneficial to creativity. Across the board, their participants strongly believed that the likely process leading to creativity is flexible thought (associative, broadly oriented thinking), rather than systematic thought (deliberate, persistent thinking). Indeed, several studies have shown that flexibility and the ability to break mental sets are related to creativity (e.g., Duncker, 1945; Smith & Blankenship, 1991). The idea that creative thinking *only* stems from flexibility is, however, too narrow. Flexibility is only one way in which people can arrive at creative ideas. There are two pathways toward creative performance: the (well known) flexibility pathway, and a *persistence* pathway (De Dreu, Baas, & Nijstad, 2008).

The *flexibility* pathway entails the ability to switch among different perspectives, involves associative thinking and requires adaptive switching among categories and approaches (Ashby, Isen, & Turken, 1999; Rowe, Hirsh, & Anderson, 2007). The flexibility pathway can, for example, be facilitated by mind wandering (Baird et al., 2012) and incubation (Dodds, Ward, & Smith, 2003; Sio & Ormerod, 2009); it is also related to positive, activating mood states, such as happiness. The *persistence* pathway, in contrast, involves effortful in-depth exploration of a few possibilities and perspectives (Chermahini & Hommel, 2010; De Dreu et al., 2008; Nijstad, De Dreu, Rietzschel, & Baas, 2010; Sagiv, Arieli, Goldenberg, & Goldschmidt, 2010). It relies on focused attention (De Dreu, Nijstad, Baas, Wolsink, & Roskes, 2012; Oberauer, Süß, Wilhelm, & Wittmann, 2008) and requires deliberate and systematic searching (Rietzschel, De Dreu, & Nijstad, 2007a). The persistence pathway can be facilitated by focused attention (De Dreu et al., 2012; Oberauer et al., 2008) and by systematic probing of a few possibilities and perspectives (Rietzschel, Nijstad, & Stroebe, 2007b), and has been found to relate to negative activating mood states, such as anger. Thus, in contrast to lay theories, which suggest that creativity is always the result of flexible thought, research has shown that multiple cognitive pathways to creativity exist.

Does Creativity Come Unexpectedly?

Laypeople tend to believe that creativity, rather than being the result of deliberate and conscious work, comes unexpectedly—for example, when traveling or with relaxing activities (Baas et al., 2015). The lay theory that creativity comes unexpectedly might, at least partially, be the result of a mystification of the creative

process by highly creative people (Van Strien, 2012). For example, Van Strien (2012) notes that many of the classic examples of sudden and unexpected insight were usually recorded years or even decades after the event supposedly took place, and do not always fit other, more contemporary accounts. Nevertheless, these stories resonate strongly with the general audience, probably because they are also in line with the previously mentioned theory about the creative individual as somehow different and beyond ordinary understanding.

Anecdotal accounts of creative individuals oftentimes report that creative discoveries resulted from a process whereby initial conscious thought is followed by an incubation phase: one is working on a problem, the problem cannot be solved, one leaves it aside for some time (i.e., the incubation period), and when returning to the problem one suddenly has some new insight into how to solve the problem.

The idea that a period of incubation might facilitate creativity has not only been suggested by lay theories and by creative people, it has also been stressed in creativity models. For example, Wallas (1926) proposed that the creative process entails four stages: preparation (acquisition of knowledge to some task, and defining the task or problem one aims to work on), incubation (unconscious task-related processing that occurs when conscious attention is diverted away from the task), illumination (a creative idea flashes into sight), and verification (the creative idea is subjected to evaluation and elaboration). Tremendous attempts have been made to scientifically investigate incubation effects. A Google Scholar search (Sio & Ormerod, 2009)—with the search restricted to the years 1997–2007 the term *incubation* along with either *creativity*, *insight*, or *problem*—yielded more than 5000 articles. Meta-analytic reviews have shown that a period of incubation indeed helps creativity (Dodds et al., 2003; Sio & Ormerod, 2009). However, it is not yet clear *why* incubation is helpful. The central discussion between different theories is about whether during an incubation period it is merely the *absence* of conscious thought that drives creativity (for example, due to relaxation, facilitating cues from the environment, forgetting of fixating elements, and mental set-shifting), or whether unconscious processes actively contribute to creative thought. Several studies provide empirical support for the idea that it is not merely the absence of conscious thought that drives creativity, but that during an incubation period unconscious processes can contribute to the generation of ideas and solutions (Ritter, van Baaren, & Dijksterhuis, 2012b; for a review, see Ritter & Dijksterhuis, 2014).

Certainly a creative idea may be found *before* a decrease in conscious effort, that is, before the incubation stage. However, during some (prolonged) creative thought processes conscious and unconscious periods alternate, and a period of incubation seems to precede creative breakthroughs.

Is Creativity Uniquely Human?

Another common lay belief is that creativity is uniquely human. Very few non-human animals are thought to have creative capabilities (behaviors that might be

interpreted as creative in humans are usually explained as instinctive or learned when it comes to nonhumans), and even if they do, these are thought to be very limited (e.g., Byrne, 1998). In fact, Guilford (1950) explicitly mentioned creativity as one of the last domains in which humans would be likely to retain superiority over machines. Many, perhaps even most, human activities could be programmed and automatized, which might raise the fear of humanity becoming obsolete somehow. Creativity might be the last stronghold of humanity: “Presumably, there would still be need for human brains to operate the machines and to invent better ones” (Guilford, 1950, p. 446). However, Artificial Intelligence (AI) can model some specific aspects of creativity, for example, transformation, exploration, and combination (Boden, 2009).

Transformational creativity entails that the space or style of an idea are transformed by altering or dropping one or more defining dimensions, allowing the generation of ideas that simply *could not* have been generated before the change.

Explorational creativity is what most creative individuals, even on the most eminent level, do—it is about exploring the spaces created by their (relatively rare) moments of transformation. Interestingly, computers can come up with exploratory processes that are comparable—sometimes even superlative—to those of highly competent human professionals (Boden, 2009, p. 27). Examples can be found in various domains, such as physics (e.g., Zytrow, 1997), music (e.g., Cope, 2006), architecture (e.g., Hersey & Freedman, 1992), and visual art (e.g., Cohen, 2002). In the visual arts, a nice example is Harold Cohen’s program, AARON (Cohen, 1995). Art made by AARON has been exhibited at major art galleries around the world. Cohen’s quote “I am a first-class colorist. But AARON is a world-class colorist” demonstrates that a computer program can surpass its programmer in creative performance.

Combinational creativity entails producing unfamiliar combinations of familiar ideas by making associations between ideas that were previously not, or only indirectly, related. For example, the creative idea of a roll-on sun cream can emerge by combining the idea of ‘sun cream’ with ‘ballpoint pen.’ The advantage of AI programs is that they can make various new combinations of familiar (already stored) concepts. What is extremely difficult for AI, however, is recognizing which combinations are valuable and, thus, useful. What is missing—as compared with the human mind—is the rich store of world knowledge and concepts (Boden, 2009). AI programs can have access to databases such as Google and, hereby, may have increased associative and inferential powers, but “using huge databases sensibly, and aptly, [...] is a tall order. Not impossible in principle, [...] but extremely difficult to achieve” (Boden, 2009, p. 26).

Thus, AI and computers can—to a certain extent—perform creatively. However, the question whether a computer could ever be ‘really’ creative is difficult to answer and may be more philosophical than psychological in nature.

Is Creativity Only About Generating Ideas?

Lay theories, as well as many scientific studies on creativity, mostly focus on the idea generation part of the creative process (West, 2002). Many popular creativity techniques, such as brainstorming (see below) focus on stimulating people's ideational output, and on ways to reduce 'blocks' that hinder people from coming up with creative ideas. However, important as idea generation is, for actual implementation of creative ideas, the most creative ideas must be recognized and selected for further development and realization. Contrary to common beliefs, the available evidence consistently demonstrates that success in idea generation does not predict success in idea selection, and that people perform at a suboptimal level (and often not better than chance) when selecting creative ideas (Rietzschel, Nijstad, & Stroebe, 2006, 2010). People tend to favor the selection of mainstream rather than creative ideas (Rietzschel et al., 2010), and research shows that even when people explicitly say that they value and endorse creativity, they can still have an implicit bias against creative ideas (Mueller, Melwani, & Goncalo, 2012). Apart from the fact that people often do not seem to value creativity, the degree to which they do also seems to vary between situations and individuals. For example, Herman and Reiter-Palmon (2011) have shown that participants with a strong *promotion focus* (i.e., a focus on growth, attaining desired outcomes, and realizing ambitions; Higgins, 1997) gave more accurate assessments of the originality of ideas. In contrast, participants with a strong *prevention focus* (i.e., a focus on safety and security, on avoiding undesirable outcomes, and fulfilling one's responsibilities) gave more accurate assessments on idea 'quality' (in this case, how coherent and 'workable' the idea was). Moreover, Mueller, Wakslak, and Krishman (2014) demonstrated that the evaluation of creative ideas can be improved by manipulating participants' construal level mindset, that is, the extent to which people's thinking is abstract or concrete (Trope & Liberman, 2010). Participants with a high-level construal (i.e., abstract) mindset rated a creative idea higher on creativity than participants with a low-level construal (i.e., concrete) mindset. Whereas these studies focused on the evaluation of ideas, De Buissonjé and colleagues (under review) went one step further—they investigated how idea selection performance can be facilitated. They have shown that idea selection performance can be enhanced by the combined effect of self-affirmation, promotion focus, and positive mood.

In sum, whereas in most creativity research the focus is on creative *idea generation*, in real-world creativity another process—*idea evaluation and selection*—is of crucial importance. Idea evaluation and selection is an essential but overlooked step in the creative process, and unless more attention is paid to this process, our understanding of creativity and innovation will remain incomplete.

Lay Theories About the Creative Place

Beside lay theories about *who* is most likely to be creative and *how* creativity works, there are several pervasive beliefs about *where* we are most likely to find creative performance; in other words, which kinds of environments and settings are most conducive to creativity. These issues have been studied particularly extensively (but not exclusively) in the field of social and especially organizational psychology. In the following, we will focus on three lay theories about the ‘creative place’: (i) the so-called three B’s (Bath, Bus, Bed) of creative ideas, (ii) productivity in brainstorming groups, and (iii) the role of freedom versus constraints in creativity.

The Three B’s of Creative Ideas

A common belief about environmental influences on creativity is that we are especially likely to get creative ideas or insights when taking a bath or a shower, when traveling, or when relaxing in general. In the creativity literature, these environmental influences are sometimes summarized as the ‘three B’s’ of creative ideas, with the three B’s referring to Bath, Bus, and Bed (Boden, 2004; Dart, 1989): places where we are in a more or less relaxed state, not actively thinking about the task or problem we were working on, and not really preoccupied with anything in particular. There are many famous anecdotes of creative discoveries made in such circumstances. One well-known example is Henri Poincaré, who experienced a major mathematical insight (which he had been searching for a long time) the moment he stepped on a bus (see, e.g., Ghiselin, 1952). Another often-described example is Friedrich von Kekulé, who discovered the ring-shaped structure of the benzene molecule while dozing by the fireside (Ghiselin, 1952). These examples appear to have contributed to several lay beliefs about creativity, such as the belief that creativity is characterized by spontaneous insights rather than deliberate thought (see our previous discussion of this issue), the notion that creativity can benefit from *incubation*, and the belief that ideas are most likely to come to us in environments where we relax, rather than work.

Most evidence for the ‘three B’s’ belief is anecdotal, and to our knowledge no systematic research has actually been done to see whether people are indeed more likely to come to creative ideas in these places, but there is research that indirectly bears on this belief. Work on the role of incubation has already been discussed in this chapter. In addition, however, the role of traveling has been studied by De Bloom, Ritter, Kühnel, Reinders, and Geurts (2014), who found that recreational travel increased participants’ flexibility (also see Gurman, 1989). Further, Maddux and Galinsky (2009) found that living abroad was associated with higher creative performance on a variety of measures, and that this relationship was mediated by the degree to which people had adapted to different cultures. Seeking out a new

environment has been suggested to work as a ‘stimulation tactic’ (Smith, 1998, a practice that can support the creative process (rather than being idea generation tools per se): exposure to a new environment can literally help people to ‘see a problem in a new light,’ or break away from habitual thoughts. Ritter et al. (2012a) demonstrated that ‘diversifying experiences’ such as unusual and unexpected events, or events that violated pre-existing schemas, stimulated participants’ cognitive flexibility. However, research by Gocłowska, Baas, Crisp, and De Dreu (2014) suggests that not everybody will be stimulated by such experiences: for people with a high need for structure, schema violations may even be detrimental for creativity.

Brainstorming and Group Creativity

In the 1950s, Alex Osborn, an advertising executive, published his book *Applied Imagination*. In this book, he described what he saw as the most common obstacle on the road to creativity: premature criticism and the resulting ‘holding back’ of ideas. To help people overcome such obstacles and generate more ideas, Osborn recommended using a procedure he called *brainstorming*. In a brainstorming session, participants are not allowed to criticize each other’s (or their own) ideas, are expected to mention every idea they can think of, no matter how weird, and are supposed to try to ‘build upon’ each other’s ideas to come up with even more creative solutions. Although brainstorming is not necessarily a group technique, Osborn recommended a group setting because of the potential for *cognitive stimulation*: by listening to other group members’ ideas, people would be stimulated to come up with new ideas they would not have thought of by themselves. Accordingly, Osborn originally predicted that “the average person can think up twice as many ideas when working with a group than when working alone” (Osborn, 1957, p. 229), and this image seems to have stuck: a large majority of people believe that group brainstorming is more effective than individual brainstorming, and people who have brainstormed in a group tend to be more satisfied with their performance than people who have worked alone (e.g., Nijstad, Stroebe, & Lodewijckx, 2006; Paulus, Dzindolet, Poletes, & Camacho, 1993; Paulus, Larey, & Ortega, 1995). In fact, brainstorming is often seen as a group technique by definition, although it is perfectly possible to brainstorm by oneself.

However, when it comes to productivity, group brainstorming is not that effective at all: groups of people brainstorming together consistently generate fewer ideas, and fewer high-quality ideas, than the same number of people working alone whose nonoverlapping ideas are pooled (so-called nominal groups). This was first demonstrated by Taylor, Berry, and Block (1958), and has since then been replicated numerous times (see Mullen, Johnson, & Salas, 1991 for an overview). Several explanations have been put forward for this productivity loss in brainstorming groups (see Stroebe, Nijstad, & Rietzschel, 2010, for an historical overview of brainstorming research), such as social loafing (people often invest less

effort in group tasks than they are capable of doing) and evaluation apprehension (people may ‘hold back’ for fear of being judged negatively by others). While such processes do seem to play a role, the strongest explanation for productivity loss is *production blocking* (Lamm & Trommsdorff, 1973), which basically comes down to cognitive interference. When brainstorming in a group, people have to take turns in expressing their ideas. If we have to wait for somebody else to stop speaking, not only can we easily forget an idea we have just come up with, but it is also difficult to continue thinking about the problem to come up with new ideas, since our cognitive resources are engaged in listening to the other person (Nijstad & Stroebe, 2006).

Nevertheless, cognitive stimulation effects have been demonstrated. For example, Dugosh, Paulus, Roland, and Yang (2000) found that individual brainstormers’ productivity was enhanced by simultaneously (while generating ideas) listening to a tape recording with another persons’ ideas. Further, Nijstad, Stroebe, and Lodewijkx (2002) found that participants generated more diverse ideas when they were presented with ‘stimulation ideas’ from a broad range of semantic categories.

The challenge, then, is to get the best of both worlds: cognitive stimulation without production blocking. There are some possibilities to achieve this. For example, the *Nominal Group Technique* (NGT; Delbecq & Van de Ven, 1971) has participants first engage in individual idea generation, followed by a sharing stage. During the sharing stage, participants are encouraged to write down and contribute any new ideas that may arise from seeing other people’s ideas. Another possibility is the use of *electronic brainstorming systems* (EBS). In an EBS, participants are seated individually behind computers, and type in their ideas individually. However, at some location on the screen, other participants’ ideas are displayed as well. Thus, both the NGT and EBS have the advantage of allowing people to generate ideas without being blocked by others, while allowing for cognitive stimulation when necessary. Research suggests that such techniques may indeed lead to productivity gains (e.g., Dennis & Valacich, 1993; Valacich, Dennis, & Connolly, 1994).

Freedom and a Lack of Constraints

Another common belief about environmental influences on creativity is the belief that creativity flourishes under circumstances of total freedom, a lack of external control, and the absence of constraints. Thus, for example, Michel de Montaigne referred to “the disposition of nature so impatient of tedious and elaborate premeditation, that if it do not go frankly and gaily to work, it can perform nothing to purpose” (De Montaigne, 1685/2012), whereas contemporary musician Lady Gaga reportedly claimed that “When you make music or write or create, it is really your job to have mind-blowing, irresponsible, condomless sex with whatever idea it is you are writing about at the time” (Goodreads.com, n.d.). What these quotes have in common is the belief that creativity requires spontaneity and freedom, rather than

control and constraints. This belief also seems to be related to the belief that young children are most creative, since they have not yet learned to focus on rewards and approval, and as such this could be considered to be at least some sort of internal consistency (Furnham, 1988) in the lay theory of creativity: freedom and an absence of constraints are considered to be essential for creativity, and young children are thought to be most free and unconstrained—and hence most creative.

Broadly speaking, this belief is largely in line with scientific findings. A substantial amount of research, again mostly done in organizational psychology, has shown that people (often: employees) are most creative when they feel autonomous and supported by their (work) environment. When people work in an environment that is supportive, nonthreatening, and challenging, they will invest more effort, are more willing to take risks (an important precondition for creativity), are more open to new ideas and opinions, and adopt a more explorative thinking style, that is, they are actively seeking out possible alternatives and improvements (e.g., Amabile, 1996; Shalley & Zhou, 2008).

Much of this work has been done from the perspective of Self-Determination Theory (SDT; e.g., Deci & Ryan, 2002). According to SDT, humans have three basic and fundamental needs: the need for autonomy, the need for competence, and the need for relatedness. Well being and motivation (particularly *intrinsic motivation*, the motivation to engage in a task for its own sake) are thought to be a function of the degree to which these basic needs are fulfilled or violated. When people perceive external control, the need for autonomy is violated, which in turn will lead to lower motivation and lower creativity (e.g., Amabile, 1996; Shalley & Perry-Smith, 2001; also see Shalley & Zhou, 2008, for an overview).

That external control can indeed kill creativity is most clearly shown in the research conducted by Teresa Amabile and colleagues (see Amabile, 1996; Hennessey & Amabile, 2010, for overviews). Most of this research has focused on the role of rewards and evaluations on creative performance. What emerges from these studies is that creativity indeed suffers when people perceive external control or pressure. Thus, for example, providing people with *controlling feedback* (e.g., telling them that their performance will be judged in order to see whether they performed as they should have) leads to lower motivation and lower creative performance than *informational feedback* (telling people that their performance will be evaluated and that they can use this evaluation to learn and to improve their future performance; Shalley & Perry-Smith, 2001). Similarly, in the domain of organizational creativity, it is generally found that contextual factors (such as organizational or team climate, leadership style, feedback and evaluation, etc.) are conducive to creativity as long as they provide support, autonomy, and challenge (Shalley & Zhou, 2008), rather than make employees feel controlled and monitored. Thus, in general, the belief that creativity requires freedom is supported by research. However, this view must be qualified in two ways.

Individual Differences

First, the general lay theory of creativity and freedom does not take into account the role of *individual differences*. Although Self-Determination Theory assumes the existence of fundamental needs that all humans share, other research has demonstrated that psychological needs differ between individuals, and that these differences can moderate the effects of contextual factors such as autonomy. Thus, for example, although job autonomy is widely considered to be an important predictor of job motivation and satisfaction, this relation has been found to depend on such individual differences as growth need strength (Hackman & Oldham, 1976), the need for autonomy and achievement (Langfred & Moye, 2004), and Personal Need for Structure (PNS; Rietzschel, Slijkhuis, & Van Yperen, 2014). In the field of creativity research, individual differences seem to make a difference as well. On the whole, autonomy contributes to creativity for those people who have the discipline and experience to work on a task independently (Chang, Huang, & Choi, 2012), who are eager to learn and feel supported to do so (Shalley, Gilson, and Blum, 2009), and who are not easily overwhelmed by a lack of structure (Rietzschel et al., 2014).

Task Complexity

Second, the problem with autonomy is that it implies complexity. That is, the more freedom people receive in how to do a task, the more they will have to figure out for themselves. The risk here is that people will respond to this cognitive load by adopting mental shortcuts that diminish complexity but may not be compatible with the demand for creativity. In this context, Ward (1994; Finke et al., 1995) formulated the *path-of-least-resistance*-hypothesis: in a creative task, people tend to generate those responses that come to mind most easily, but unfortunately those often are the least creative responses. For example, when asked to generate creative ideas as to how people can improve their health, the first things to come to mind are the suggestions all of us are confronted with every day (e.g., eat more vegetables, stop smoking). Since creativity requires ideas, solutions, or products that are novel, original, or unusual, anything that makes it difficult to leave the path of least resistance can be a hindrance to creative performance. Research suggests that task complexity can indeed increase reliance on mental heuristics (e.g., Bodenhausen & Lichtenstein, 1987; Branscombe & Cohen, 1991; Ford & Kruglanski, 2005; Simon, 1955; Van Prooijen & Van de Veer, 2010), probably because complex tasks put a heavy load on working memory (WM), especially the *central executive* component of WM (Baddeley, 1996). Since WM capacity has been linked repeatedly to creative performance (De Dreu et al., 2012; Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; Lee & Theriault, 2013), it seems plausible that the complexity associated with high autonomy can inhibit creative performance through a reliance on mental shortcuts, such as the path of least resistance.

Lay Theories About the Creative Product

Creative ideas and products can be extremely varied, from musical masterpieces, to paintings, to literary work, to scientific and technological breakthroughs, and to creative solutions for problems. Often, lay beliefs about creativity suggest that the term creativity only applies to revolutionary ideas, for example, ideas we give Pulitzer and Nobel prizes for, and not to ideas that enhance and enrich our lives, for example, creating a new recipe. As described earlier (see ‘Lay theories about the creative person’), most creativity experts, however, rely on a continuity principle and assume that highly eminent creativity (e.g., Big-C creativity) is *not* fundamentally different from ordinary creativity (e.g., little-c creativity). This implies that revolutionary ideas as well as ideas that enhance and enrich our lives can be considered creative—what differs is the level of creativity.

Is ‘Creativity’ Always Subjective?

Irrespective of the level of creativity, the question rises whether the creativity of an idea or product can be evaluated with sufficient reliability and validity, or whether it merely depends on ‘the eye of the beholder.’ Indeed, we expect that most creativity researchers share our experience of having people, both in lay or student audiences and among noncreativity researchers, coming up to us and asking somewhat sceptically how it is possible to measure creative performance, ‘since it is all subjective, anyway.’ Is it possible to measure creativity at all? Whereas some researchers see the evaluation of creativity as inherently subjective (e.g., Kilgour, Sasser, & Koslow, 2013), or even as depending on zeitgeist (i.e., the sociocultural environment an idea or product is born into; Simonton, 1999), others assume that within a certain time and group, people tend to agree on whether an idea or product can be considered more or less creative (e.g., Amabile, 1982; Guilford, 1967; Lim & Plucker, 2001; Runco, 1999; Runco & Johnson, 2002). The judgment of creativity is, to a certain degree, subjective; this subjectivity, however, does not have to be problematic, as long as different people get to (more or less) the same subjective judgment. Therefore, instead of striving for ‘objectivity’ in creativity judgment, we should strive for *consensus* in creativity judgment. The challenge is to identify relevant criteria of an idea’s or product’s creativity.

Creativity researchers have tried to define the characteristics that lead to an idea or product being judged as ‘creative.’ There is strong agreement (also see our earlier explanation on the definition of creativity) that a creative idea or product has to be original (i.e., novel) and useful (i.e., effective; e.g., Amabile, 1996; Campbell, 1960; Feldman, Csikszentmihalyi, & Gardener, 1994; Runco, 2004). As we have seen, these two criteria are part of the commonly used definition of creativity, but the question is whether people also use these two dimensions to judge whether something is creative. Empirical research supports this idea by showing that

creativity evaluations strongly depend on the perceived novelty, and, to a lesser degree, on the perceived usefulness (Caroff & Besançon, 2008; Dietrich & Haider, 2015; Runco & Charles, 1993).

The evaluation of an idea or product on specific criteria usually entails asking judges to evaluate the idea or product on a dichotomous or continuous scale (Benedek, Mühlmann, Jauk, & Neubauer, 2013; Kaufman, Plucker, & Baer, 2008; Silvia et al., 2008). This method is rooted in the *Consensual Assessment Technique* (CAT; Amabile, 1982; Hennessey & Amabile, 1999)—by far the most common method in creativity research, due to its relative simplicity and the consistently high levels of inter-rater reliability for various kinds of creative products. The assessment of ideas or products based on the CAT is particularly useful in the study of ‘little-c’ (everyday) creativity, and it can be applied to the creativity evaluation of any kind of idea or product.

Using the CAT technique, the ideas generated by participants are generally scored for fluency, flexibility, elaboration, and originality. Fluency is operationalized as the number of ideas generated. Flexibility is operationalized as the number of unique categories the ideas can be assigned to. Elaboration is operationalized as the amount of detail that is provided (Guilford, 1968). Originality is operationalized as the uniqueness of the idea generated (Runco, 1999).

Consequences of Lay Theories on Creativity

In this final section, we will focus on the possible negative consequences of incorrect lay theories of creativity. Why is it a problem if people hold false or incomplete beliefs about creative persons, processes, places, or products? We will discuss three reasons: stereotype threat, lack of developmental opportunities, and self-selection.

Stereotype Threat

Some of the lay theories we have discussed take the form of stereotypes (e.g., about older people). Although stereotypes need not be negative in content (for example, some social groups may be stereotyped as ‘smart’ or ‘very trustworthy’), they have been linked extensively to prejudice and discrimination (e.g., Bar-Tal, Graumann, Kruglanski, & Stroebe, 2013). Another way in which stereotypes can be problematic is by eliciting so-called *stereotype threat* (Steele & Aronson, 1995). When members of a certain social group are aware of the existing stereotype of their group, the fear of confirming this stereotype may cause them to underperform in exactly the kinds of situations the stereotype relates to. Thus, stereotype threat can cause stereotypes to become a self-fulfilling prophecy.

Stereotype threat effects have also been demonstrated in creativity research. For example, Seibt and Förster (2004, Study 4) found that participants performed worse on a creative idea generation task when they had first been confronted with a negative stereotype concerning their own group (i.e., that students from their major typically performed badly on these tasks) than when they had been confronted with a positive stereotype. Thus, some people's creative performance might suffer if they do not fit the stereotype of the creative individual, but rather are stereotyped as dull and unimaginative. For example, people who clearly do *not* have the traits commonly associated with creativity (e.g., people who are emotionally stable, not impulsive, highly systematic, not flexible, etc.) might a priori be seen as uncreative and might be exposed to such stereotypes, thus causing stereotype threat and lower creative performance. This could prevent them from realizing or developing the creative potential they may actually have (e.g., because they might be well suited to performing creatively in a more systematic fashion) and, over time, confirm their self-image of not being creative.

Lack of Developmental Opportunities

If people hold incorrect beliefs about the abilities and processes underlying creativity, or about the contextual factors that stimulate or inhibit creativity, interventions to stimulate creativity may be less effective or even fail completely. For example, creativity trainings can be effective, but how effective they are depends on the type of training offered. In a meta-analysis of studies on creativity training, Scott et al. (2004) concluded that “successful training courses devote less time and resources to techniques that stress unconstrained exploration” (p. 377). Instead, trainings that provide people with concrete and specific techniques and heuristics (such as the use of brainstorming rules, checklists, or feature comparisons) appear to be significantly more effective. The relation with creativity beliefs lies in the fact that, as we have seen, creativity is often particularly associated with unconstrained exploration, rather than with systematic work. This could easily lead people to prefer training methods or content (or other interventions) that fit this association, even though such trainings tend to be less effective.

A more general problem in this context is that, as also addressed earlier in this chapter, some lay beliefs about creativity seem to imply an *entity theory* (Dweck et al., 1993) of creative potential, that is, the belief that people are either creative or not, and that creativity cannot be developed (e.g., Furnham, 2014). Research shows that entity theories (as compared with *incremental* theories) can be quite detrimental for learning and performance in a variety of settings and domains (see Burnette et al., 2013, for a meta-analysis). For example, Plaks and Chasteen (2013) found that older adults performed worse on a variety of memory tasks if they had entity beliefs than if they had incremental beliefs. In a study among math teachers, Rattan, Good, and Dweck (2012) found that teachers holding entity theories tended to use feedback strategies that were comforting (e.g., reassuring underperforming students

that ‘not everybody can be good at maths’), but also demotivating. As we have seen, creative skills can be developed (Scott et al., 2004), but if people hold entity theories about creativity, they are less likely to seek out such development opportunities (or to offer such opportunities to others, e.g., in the case of supervisors) and are less likely to attain optimal learning outcomes.

Self-selection

If creativity is consistently associated with certain traits or environments, some people may opt out of the creative process, or certain environments, because of a perceived lack of fit. When it comes to work, for example, the attraction-selection-attrition (ASA) framework (Schneider, 1987) describes how only a subset of people will be *attracted* to certain jobs or organizations, how only a subset of these people will actually be *selected* for a job within the organization, and, finally, how only a subset of those people will *remain* within the organization over time. One consequence of this is that organizations run the risk of becoming more homogeneous over time, and hence losing out on potentially valuable diversity. Thus, some people might not feel attracted to organizations with a reputation for creativity or innovation, or to professions that are known (or thought) to require creativity. If such self-selection happens on the basis of misconceptions, both the organization and the individual might be worse off.

Conclusion

In this chapter, we have attempted to outline and discuss several of the lay theories that people hold about creativity. Some of these lay theories concern the characteristics of creative persons, such as the ‘mad genius’ belief, others revolve around the creative process itself (such as the ‘flexibility’ belief) or the environmental factors that are thought to contribute to (or hinder) creative performance (such as group collaboration). As we have seen, these lay theories are not always in line with scientific findings; many are completely false, and even the ones that have a basis in fact are only *partially* correct. This is important, because misconceptions and incorrect (or incomplete) assumptions about creativity can be harmful, for a variety of reasons (such as stereotype threat and a lack of development opportunities). Given the importance of creativity for all domains of life, including such diverse endeavors as science, technology, design, sports, medicine, and art, we cannot afford to let our lay theories guide our creative efforts without empirical scrutiny.

References

- Amabile, T. M. (1982). Children's artistic creativity: Detrimental effects of competition in a field setting. *Personality and Social Psychology Bulletin*, 8, 573–578.
- Amabile, T. M. (1983). *The social psychology of creativity*. New York: Springer.
- Amabile, T. M. (1996). *Creativity in context: Update to "the social psychology of creativity"*. Boulder, CO: Westview Press.
- Ashby, F. G., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, 106, 529–550.
- Baas, M., Koch, S., Nijstad, B. A., & De Dreu, C. K. (2015). Conceiving creativity: The nature and consequences of laypeople's beliefs about the realization of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 9, 340.
- Baas, M., Nijstad, B. A., Boot, N. C., & De Dreu, C. K. W. (2016). Mad genius revisited: Vulnerability to psychopathology, biobehavioral approach-avoidance, and creativity. *Psychological Bulletin*, 142, 668–692.
- Bachelor, P., & Michael, W. B. (1991). Higher-order factors of creativity within Guilford's structure-of-intellect model: A re-analysis of a fifty-three variable data base. *Creativity Research Journal*, 4, 157–175.
- Bachelor, P. A., & Michael, W. B. (1997). The structure of intellect model revisited. *The Creativity Research Handbook*, 1, 155–182.
- Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, 49, 5–28.
- Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological Science*, 23, 1117–1122.
- Barron, F. (1955). The disposition toward originality. *The Journal of Abnormal and Social Psychology*, 51, 478–485.
- Bar-Tal, D., Graumann, C. F., Kruglanski, A. W., & Stroebe, W. (Eds.). (2013). *Stereotyping and prejudice: Changing conceptions*. New York, NY: Springer Science & Business Media.
- Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions on intelligence and creativity. *Intelligence*, 46, 73–83.
- Benedek, M., Mühlmann, C., Jauk, E., & Neubauer, A. C. (2013). Assessment of divergent thinking by means of the subjective top-scoring method: Effects of the number of top-ideas and time-on-task on reliability and validity. *Psychology of Aesthetics, Creativity, and the Arts*, 7, 341–349.
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. London, UK: Routledge.
- Boden, M. A. (2009). Computer models of creativity. *AI Magazine*, 30, 23–24.
- Bodenhausen, G. V., & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: The impact of task complexity. *Journal of Personality and Social Psychology*, 52, 871–880.
- Branscombe, N. R., & Cohen, B. M. (1991). Motivation and complexity levels as determinants of heuristic use in social judgment. In J. Forgas (Ed.), *Emotion and social judgments* (pp. 145–160). Oxford, England: Pergamon Press.
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139, 655–701.
- Byrne, R. (1998). The early evolution of creative thinking: Evidence from monkeys and apes. In S. Mithen (Ed.), *Creativity in human evolution and prehistory* (pp. 110–124). London, UK: Routledge.
- Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, 67, 380–400.

- Caroff, X., & Besançon, M. (2008). Variability of creativity judgments. *Learning and Individual Differences, 18*, 367–371.
- Carver, C. S., Sutton, S. K., & Scheier, M. F. (2000). Action, emotion, and personality: Emerging conceptual integration. *Personality and Social Psychology Bulletin, 26*, 741–751.
- Chang, J. W., Huang, D. W., & Choi, J. N. (2012). Is task autonomy beneficial for creativity? Prior task experience and self-control as boundary conditions. *Social Behavior and Personality, 40*, 705–724.
- Chermahini, S. A., & Hommel, B. (2010). The (b)link between creativity and dopamine: Spontaneous eye blink rates predict and dissociate divergent and convergent thinking. *Cognition, 115*, 458–465.
- Chomsky, C. (1972). Stages in language development and reading exposure. *Harvard Educational Review, 42*, 1–33.
- Cohen, H. (1995). The further exploits of AARON Painter. In S. Franchi & G. Guzeldere (Eds.), *Constructions of the mind: Artificial intelligence and the humanities* (Vol. 4, pp. 141–160). Stanford, CA: Special edition of Stanford Humanities Review.
- Cohen, H. (2002). A Million Millennial Medicis. In L. Candy & E. A. Edmonds (Eds.), *Explorations in art and technology* (pp. 91–104). London, UK: Springer.
- Cope, D. (2006). *Computer models of musical creativity*. Cambridge, MA: The MIT Press.
- Creating Minds. (n.d.). *Age and creativity*. Retrieved on August 4th, 2016, from <http://creatingminds.org/articles/age.htm>
- Cropley, A. J. (1990). Creativity and mental health in everyday life. *Creativity Research Journal, 3*, 167–178.
- Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. New York, NY: Basic Books.
- Cuddy, A. J. C., Norton, M. I., & Fiske, S. T. (2005). This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues, 61*, 267–285.
- Dart, P. (1989). *Bus-bath-bed: A rationale for irrational predicate identifications in the service of creativity*. Paper presented at the Annual Meeting of the Speech Communication Association (75th, San Francisco, CA, November 18–21, 1989).
- De Bloom, J., Ritter, S., Kühnel, J., Reinders, J., & Geurts, S. (2014). Vacation from work: A 'ticket to creativity'? The effects of recreational travel on cognitive flexibility and originality. *Tourism Management, 44*, 164–171.
- De Buissonjé, D. R., Ritter, S. M., De Bruin, S., Ter Horst, J. M.-L., & Meeldijk, A. (under review). *Facilitating creative idea selection: The combined effects of self-affirmation, promotion focus and positive affect*. Manuscript submitted for publication.
- Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- De Dreu, C. K., Baas, M., & Nijstad, B. A. (2008). Hedonic tone and activation level in the mood-creativity link: Toward a dual pathway to creativity model. *Journal of Personality and Social Psychology, 94*, 739–756.
- De Dreu, C. K., Nijstad, B. A., Baas, M., Wolsink, I., & Roskes, M. (2012). Working memory benefits creative insight, musical improvisation, and original ideation through maintained task-focused attention. *Personality and Social Psychology Bulletin, 38*, 656–669.
- de Montaigne, M. (1685/1877/2012). *Essays of Michel de Montaigne*. Available <http://www.gutenberg.org/files/3600/3600-h/3600-h.htm> [April 26, 2016].
- Delbecq, A. L., & Van de Ven, A. H. (1971). A group process model for problem identification and program planning. *The Journal of Applied Behavioral Science, 7*, 466–492.
- Dennis, A. R., & Valacich, J. S. (1993). Computer brainstorming: More heads are better than one. *Journal of Applied Psychology, 78*, 531–537.
- Dietrich, A., & Haider, H. (2015). Human creativity, evolutionary algorithms, and predictive representations: The mechanics of thought trials. *Psychonomic Bulletin & Review, 22*, 897–915. doi:10.3758/s13423-014-0743-x.
- Dodds, R. A., Ward, T. B., & Smith, S. M. (2003). A review of the experimental literature on incubation in problem solving and creativity. *Creativity Research Handbook, 3*, 285–302.

- Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yang, H. C. (2000). Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology, 79*, 722–735.
- Duncker, K. (1945). On problem-solving. *Psychological Monographs, 58*(5), i-113.
- Dweck, C. S., Hong, Y., & Chiu, C. (1993). Implicit theories individual differences in the likelihood and meaning of dispositional inference. *Personality and Social Psychology Bulletin, 19*, 644–656.
- Fasko, D. (2001). Education and creativity. *Creativity Research Journal, 13*, 317–327.
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2*, 290–309.
- Feist, G. J., & Gorman, M. E. (1998). The psychology of science: Review and integration of a nascent discipline. *Review of General Psychology, 2*, 3–47.
- Feldman, D. H., Csikszentmihalyi, M., & Gardner, H. (1994). *Changing the world: A framework for the study of creativity*. Greenwich, CT: Greenwood Publishing Group.
- Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.
- Finke, R. A., Ward, T. B., & Smith, S. M. (1995). *The Creative cognition approach*. Boston, MA: MIT press.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902.
- Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*, 1011.
- Florida, R. (2002). *The rise of the creative class. And how it's transforming work, leisure and everyday life*. New York, NY: Basic Books.
- Ford, T. E., & Kuglanski, A. W. (2005). Effects of epistemic motivations on the use of accessible constructs in social judgment. *Personality and Social Psychology Bulletin, 21*, 950–962.
- Friedman, R. S., & Förster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology, 81*, 1001–1013.
- Furnham, A. (2014). Increasing your intelligence: Entity and incremental beliefs about the multiple “intelligences”. *Learning and Individual Differences, 32*, 163–167.
- Furnham, A. F. (1988). *Lay theories: Everyday understanding of problems in the social sciences*. Oxford, UK: Pergamon Press.
- Gardner, H. (1993). *Creating minds*. New York: BasicBooks.
- Ghiselin, B. (1952). *The creative process*. Berkeley, CA: University of California Press.
- Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J.-P. (2008). The effects of aging on researchers' publication and citation patterns. *PLoS ONE, 3*, e4048.
- Gocłowska, M. A., Baas, M., Crisp, R. J., & De Dreu, C. K. (2014). Whether social schema violations help or hurt creativity depends on need for structure. *Personality and Social Psychology Bulletin, 40*, 959–971.
- Goodreads.com. (n.d.). *Lady Gaga quotes*. Retrieved from <http://www.goodreads.com/quotes/201754-when-you-make-music-or-write-or-create-it-s-really>
- Greene, G. (1980). *Ways of escape*. New York, NY: Simon & Schuster.
- Gregoire, C. (2016, March 4). 18 Things highly creative people do differently. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/2014/03/04/creativity-habits_n_4859769.html
- Gregoire, C., & Kaufman, S. B. (2016, January 4). Creative people's brains really do work differently. *Quartz*. Retrieved from <http://qz.com/584850/creative-peoples-brains-really-do-work-differently/>
- Griskevicius, V., Cialdini, R. B., & Kenrick, D. T. (2006). Peacocks, Picasso, and parental investment: The effects of romantic motives on creativity. *Journal of Personality and Social Psychology, 91*, 63–76.
- Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444–454.
- Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior, 1*, 3–14.

- Guilford, J. P. (1968). *Intelligence, creativity, and their educational implications*. San Diego, CA: RR Knapp.
- Gurman, E. B. (1989). Travel abroad: A way to increase creativity? *Educational Research Quarterly*, *13*, 12–16.
- Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, *16*, 250–279.
- Hennessey, B. A., & Amabile, T. M. (1999). Consensual assessment. *Encyclopedia of Creativity*, *1*, 347–359.
- Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology*, *61*, 569–598.
- Herman, A., & Reiter-Palmon, R. (2011). The effect of regulatory focus on idea generation and idea evaluation. *Psychology of Aesthetics, Creativity, and the Arts*, *5*, 13–20.
- Hersey, G., & Freedman, R. (1992). *Possible Palladian villas*. Cambridge: Massachusetts MIT.
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.
- Hirt, E. R., Devers, E. E., & McCrea, S. M. (2008). I want to be creative: Exploring the role of hedonic contingency theory in the positive mood-cognitive flexibility link. *Journal of Personality and Social Psychology*, *94*, 214–230.
- Howkins, J. (2002). *The creative economy: How people make money from ideas*. London, UK: Penguin.
- Kaiser, K. (n.d.). *20 things only highly creative people would understand*. Retrieved from <http://www.lifehack.org/articles/communication/20-things-remember-you-love-highly-creative-person.html>
- Kaufman, J. C. (2002). Dissecting the golden goose: Components of studying creative writers. *Communication Research Journal*, *14*, 27–40.
- Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four C model of creativity. *Review of General Psychology*, *13*, 1–12.
- Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment* (Vol. 53). Hoboken, NJ: Wiley.
- Kilgour, M., Sasser, S., & Koslow, S. (2013). Creativity awards: Great expectations? *Creativity Research Journal*, *25*, 163–171.
- Lamm, H., & Trommsdorff, G. (1973). Group versus individual performance on tasks requiring ideational proficiency (brainstorming): A review. *European Journal of Social Psychology*, *3*, 361–388.
- Lamont, R. A., Swift, H. J., & Abrams, D. (2015). A review and meta-analysis of age-based stereotype threat: Negative stereotypes, not facts, do the damage. *Psychology and Aging*, *30*, 180–193.
- Langfred, C. L., & Moye, N. A. (2004). Effects of task autonomy on performance: An extended model considering motivational, informational, and structural mechanisms. *Journal of Applied Psychology*, *89*, 934–945.
- Lee, C. S., & Theriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, *41*, 306–320.
- Lim, W., & Plucker, J. A. (2001). Creativity through a lens of social responsibility: Implicit theories of creativity with Korean samples. *The Journal of Creative Behavior*, *35*, 115–130.
- Mackinnon, D. W. (1962). The nature and nurture of creative talent. *American Psychologist*, *17*, 484–495.
- Maddux, W. W., & Galinsky, A. D. (2009). Cultural borders and mental barriers: The relationship between living abroad and creativity. *Journal of Personality and Social Psychology*, *96*, 1047–1061.
- Maier, N. R. (1967). Assets and liabilities in group problem solving: The need for an integrative function. *Psychological Review*, *74*, 239–294.
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220–232.
- Mithen, S. J. (1998). *Creativity in human evolution and prehistory*. London, UK: Routledge.

- Mueller, J. S., Melwani, S., & Goncalo, J. A. (2012). The bias against creativity why people desire but reject creative ideas. *Psychological Science*, 23, 13–17.
- Mueller, J. S., Wakslak, C. J., & Krishnan, V. (2014). Construing creativity: The how and why of recognizing creative ideas. *Journal of Experimental Social Psychology*, 51, 81–87.
- Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology*, 12, 3–23.
- Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creativity Research Journal*, 15, 107–120.
- Mumford, M. D., Marks, M. A., Connelly, M. S., Zaccaro, S. J., & Johnson, J. F. (1998). Domain-based scoring in divergent-thinking tests: Validation evidence in an occupational sample. *Creativity Research Journal*, 11, 151–163.
- Ng, T. W. H., & Feldman, D. C. (2008). The relationship of age to ten dimensions of job performance. *Journal of Applied Psychology*, 93, 392–423.
- Ng, T. W. H., & Feldman, D. C. (2012). Evaluating six common stereotypes about older workers with meta-analytical data. *Personnel Psychology*, 65, 821–858.
- Ng, T. W. H., & Feldman, D. C. (2013). A meta-analysis of the relationships of age and tenure with innovation-related behaviour. *Journal of Occupational and Organizational Psychology*, 86, 585–616.
- Nickerson, R. S. (1999). Enhancing Creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 392–430). Cambridge, UK: Cambridge University Press.
- Nielsen, J. A., Zielinski, B. A., Fletcher, P. T., Alexander, A. L., Lange, N., Bigler, E. D., et al. (2014). Abnormal lateralization of functional connectivity between language and default mode regions in autism. *Molecular Autism*, 5, 1–8.
- Nijstad, B. A., De Dreu, C. K., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, 21, 34–77.
- Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: A cognitive model of idea generation in groups. *Personality and Social Psychology Review*, 10, 186–213.
- Nijstad, B. A., Stroebe, W., & Lodewijckx, H. F. M. (2002). Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology*, 38, 535–544.
- Nijstad, B. A., Stroebe, W., & Lodewijckx, H. F. M. (2006). The illusion of group productivity: A reduction of failures explanation. *European Journal of Social Psychology*, 36, 31–48.
- Oberauer, K., Süß, H. M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence*, 36, 641–652.
- Osborn, A. F. (1957). *Applied imagination* (2nd ed.). New York, NY: Charles Scribner's Sons.
- Paulus, P. B., Dzindolet, M. T., Poletes, G., & Camacho, L. M. (1993). Perception of performance in group brainstorming: The illusion of group productivity. *Personality and Social Psychology Bulletin*, 19, 78–89.
- Paulus, P. B., Larey, T. S., & Ortega, A. H. (1995). Performance and perception of brainstormers in an organizational setting. *Basic and Applied Social Psychology*, 17, 249–265.
- Pinker, S. (1984). Visual cognition: An introduction. *Cognition*, 18, 1–63.
- Plaks, J. E., & Chasteen, A. L. (2013). Entity versus incremental theories predict older adults' memory performance. *Psychology and Aging*, 28, 948–957.
- Plucker, J. A., & Renzulli, J. S. (1999). Psychometric approaches to the study of human creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 35–61). Cambridge, UK: Cambridge University Press.
- Rattan, A., Good, C., & Dweck, C. S. (2012). “It’s ok—Not everyone can be good at math”: Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, 48, 731–737.
- Reiter-Palmon, R., Mumford, M. D., & Threlfall, K. V. (1998). Solving everyday problems creatively: The role of problem construction and personality type. *Creativity Research Journal*, 11, 187–197.
- Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan*, 42, 305–310.

- Rietzschel, E. F., De Dreu, C. K. W., & Nijstad, B. A. (2007a). Personal need for structure and creative performance: The moderating influence of fear of invalidity. *Personality and Social Psychology Bulletin*, *33*, 855–866.
- Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2006). Productivity is not enough: A comparison of interactive and nominal brainstorming groups on idea generation and selection. *Journal of Experimental Social Psychology*, *42*, 244–251.
- Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2007b). Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and originality of generated ideas. *Journal of Experimental Social Psychology*, *43*, 933–946.
- Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2010). The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British Journal of Psychology*, *101*, 47–68.
- Rietzschel, E. F., Slijkhuis, J. M., & Van Yperen, N. W. (2014). Close monitoring as a contextual stimulator: How need for structure affects the relation between close monitoring and work outcomes. *European Journal of Work and Organizational Psychology*, *23*, 394–404.
- Rietzschel, E. F., Zacher, H., & Stroebe, W. (2016). A lifespan perspective on creativity and innovation at work. *Work, Aging and Retirement*, *2*, 105–129.
- Ritter, S. M., Damian, R. I., Simonton, D. K., van Baaren, R. B., Strick, M., Derks, J., et al. (2012a). Diversifying experiences enhance cognitive flexibility. *Journal of Experimental Social Psychology*, *48*, 961–964.
- Ritter, S. M., & Dijksterhuis, A. (2014). The unconscious foundations of creative thought. *Frontiers in Human Neuroscience*, *8*, 1–10.
- Ritter, S. M., van Baaren, R. B., & Dijksterhuis, A. (2012b). Creativity: The role of unconscious processes in idea generation and idea selection. *Thinking Skills and Creativity*, *7*, 21–27.
- Ritter, S. M., & Mostert, N. (2016). Enhancement of creative thinking skills using a cognitive based creativity training. *Journal of Cognitive Enhancement*.
- Ritter, S. M., Strick, M., Bos, M. W., van Baaren, R. B., & Dijksterhuis, A. (2012c). Good morning creativity: Task reactivation during sleep enhances beneficial effect of sleep on creative performance. *Journal of Sleep Research*, *21*, 643–647.
- Roskes, M., De Dreu, C. K. W., & Nijstad, B. A. (2012). Necessity is the mother of invention: Avoidance motivation stimulates creativity through cognitive effort. *Journal of Personality and Social Psychology*, *103*, 242–256.
- Rowe, G., Hirsh, J. B., & Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences*, *104*, 383–388.
- Runco, M. A. (1988). Creativity research: Originality, utility, and integration. *Creativity Research Journal*, *1*, 1–7.
- Runco, M. A. (1999). A longitudinal study of exceptional giftedness and creativity. *Creativity Research Journal*, *12*, 161–164.
- Runco, M. A. (2004). Personal creativity and culture. In S. Lau, A. N. B. Hui, & G. Y. C. Ng (Eds.), *Creativity when east meets west* (pp. 9–22). Singapore: World Scientific Publishing.
- Runco, M. A., & Charles, R. E. (1993). Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences*, *15*, 537–546.
- Runco, M. A., & Johnson, D. J. (2002). Parents' and teachers' implicit theories of children's creativity: A cross-cultural perspective. *Creativity Research Journal*, *14*, 427–438.
- Runco, M. A., Johnson, D. J., & Bear, P. K. (1993). Parents' and teachers' implicit theories of children's creativity. *Child Study Journal*.
- Sagiv, L., Arieli, S., Goldenberg, J., & Goldschmidt, A. (2010). Structure and freedom in creativity: The interplay between externally imposed structure and personal cognitive style. *Journal of Organizational Behavior*, *31*, 1086–1110.
- Schneider, B. (1987). The people make the place. *Personnel Psychology*, *40*, 437–453.
- Scott, G., Leritz, L. E., & Mumford, M. D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal*, *16*, 361–388.
- Scratchley, L. S., & Hakstian, A. R. (2001). The measurement and prediction of managerial creativity. *Creativity Research Journal*, *13*, 367–384.

- Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory focus. *Journal of Personality and Social Psychology, 87*, 38–56.
- Shalley, C. E., Gilson, L. L., & Blum, T. C. (2009). Interactive effects of growth need strength, work context, and job complexity on self-reported creative performance. *Academy of Management Journal, 52*, 489–505.
- Shalley, C. E., & Perry-Smith, J. E. (2001). Effects of social-psychological factors on creative performance: The role of informational and controlling expected evaluation and modeling experience. *Organizational Behavior and Human Decision Processes, 84*, 1–22.
- Shalley, C. E., & Zhou, J. (2008). Organizational creativity research: A historical overview. In C. Shalley & J. Zhou (Eds.), *Handbook of organizational creativity* (pp. 3–31). Mahwah, NJ: Lawrence Erlbaum.
- Shearing, H. A. (1992). Creativity and older adults. *Leadership & Organization Development Journal, 13*, 11–16.
- Silvia, P. J., & Kaufman, J. C. (2010). Creativity and mental illness. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 381–394). Cambridge, UK: Cambridge University Press.
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., et al. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts, 2*, 68–85.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99–118.
- Simon, H. A., Newell, A., & Shaw, J. C. (1962). The processes of creative thinking. In H. E. Gruber, G. Terrell, & M. Wertheimer (Eds.), *Contemporary approaches to creative thinking* (pp. 63–119). New York, NY: Lieber-Atherton.
- Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104*, 66–89.
- Simonton, D. K. (1999). Creativity as blind variation and selective retention: Is the creative process Darwinian? *Psychological Inquiry, 10*, 309–328.
- Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge, UK: Cambridge University Press.
- Simonton, D. K. (2014a). The mad-genius paradox: Can creative people be more mentally healthy but highly creative people more mentally ill? *Perspectives on Psychological Science, 9*, 470–480.
- Simonton, D. K. (2014b). More method in the mad-genius controversy: A historiometric study of 204 historic creators. *Psychology of Aesthetics, Creativity, and the Arts, 8*, 53–61.
- Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135*, 94–120.
- Sligte, D. J., De Dreu, C. K., & Nijstad, B. A. (2011). Power, stability of power, and creativity. *Journal of Experimental Social Psychology, 47*, 891–897.
- Smith, G. F. (1998). Idea-generation techniques: A formulary of active ingredients. *The Journal of Creative Behavior, 32*, 107–134.
- Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *The American Journal of Psychology, 104*, 61–87.
- Smith, S. M., Ward, T. B., & Finke, R. A. (1995). *The creative cognition approach*. Cambridge, MA: MIT Press.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811.
- Stein, M. I. (1953). Creativity and culture. *The Journal of Psychology, 36*, 311–322.
- Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. *Handbook of creativity, 1*, 3–15.
- Sternberg, R. J., & O'Hara, L. A. (1999). Creativity and intelligence. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 251–272). Cambridge, MA: Cambridge University Press.
- Stroebe, W. (2010). Age and scientific creativity. Wiley, Inc.

- Stroebe, W., Nijstad, B. A., & Rietzschel, E. F. (2010). Beyond productivity loss in brainstorming groups: The evolution of a question. *Advances in Experimental Social Psychology*, *43*, 157–203. Burlington, MA: Academic Press.
- Taylor, D. W., Berry, P. C., & Block, C. H. (1958). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly*, *3*, 23–47.
- Treffinger, D. J. (1995). Creative problem solving: Overview and educational implications. *Educational Psychology Review*, *7*, 301–312.
- Trope, Yaacov, & Liberman, Nira. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463.
- Valacich, J. S., Dennis, A. R., & Connolly, T. (1994). Idea-generation in computer-based groups: A new ending to an old story. *Organizational Behavior and Human Decision Processes*, *57*, 448–467.
- Van Prooijen, J.-W., & Van De Veer, E. (2010). Perceiving pure evil: The influence of cognitive load and prototypical evilness on demonizing. *Social Justice Research*, *23*, 259–271.
- Van Strien, P. J. (2012). *Psychologie van de wetenschap (Psychology of science)*. Amsterdam, The Netherlands: Amsterdam University Press.
- Van Strien, P. J. (2015). *Het creatieve genie (The creative genius)*. Amsterdam, The Netherlands: Amsterdam University Press.
- Vincent, A. S., Decker, B. P., & Mumford, M. D. (2002). Divergent thinking, intelligence, and expertise: A test of alternative models. *Creativity Research Journal*, *14*, 163–178.
- Wallas, G. (1926). *The art of thought*. New York: Harcourt Brace.
- Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, *27*, 1–40.
- Ward, T. B., Finke, R. A., & Smith, S. M. (1995). *Creativity and the mind*. New York, NY: Plenum.
- West, M. A. (2002). Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups. *Applied Psychology*, *51*, 355–387.
- Zytkow, J. (1997). *Machine discovery*. London, UK: Kluwer Academic.

Mindsets About Malleability and Intergroup Relations

Aneeta Rattan and Oriane Georgeac

Stereotyping. Prejudice. Discrimination. We live in a world rife with unwanted intergroup bias. Is this inevitable, or can it be changed? Recent research suggests that people's yes or no answers to this question may determine which reality emerges, one in which intergroup relations are improved over time, or one in which they are continually marked by intergroup divisions and bias. That is, a burgeoning field of research shows that people's ideas about whether attributes can change or not—their naïve beliefs about malleability—have real consequences for intergroup relations.

The main goal of this chapter is to review the literature on these *lay theories* or *mindsets* about malleability (terms that will be used interchangeably) to illustrate how people's mindsets drive their outlook on and responses to stereotyping, prejudice, discrimination, and the likelihood of intergroup reconciliation. The central tenet of this chapter is that a lay theory approach offers much to the study of intergroup relations. Reciprocally, of course, the unique challenges of intergroup interactions also offer novel insights to the study of lay theories about malleability. Therefore, along the way, we will take opportunities to highlight some of the many open questions that may benefit from integrating the study of mindsets about malleability and intergroup relations.

To ground our discussion in precise psychological terms, we offer definitions of stereotyping (Cardwell, 1996), prejudice (Allport, 1954), discrimination (Dovidio, Hewstone, Glick, & Esses, 2013), and stigma (Goffman, 1963).

Stereotyping: The cognitive association of social groups with specific, positive or negative, traits or characteristics.

Prejudice: Affective negativity toward outgroups, which can be directed toward outgroup members on the basis of their group memberships.

A. Rattan (✉) · O. Georgeac
Organisational Behaviour, London Business School, Regent's Park,
Sussex Place, London NW1 4SA, UK
e-mail: arattan@london.edu

Discrimination: The differential and negative treatment of a person or group of people, due to their social group memberships. It can be a result of conscious or nonconscious stereotyping, prejudice, or ingroup preference.

Stigma: A characteristic that marks a person as “lesser than” in the minds of others, or the experience of being treated as “less than” due to the possession of said characteristic.

Although these cognitive, affective, and behavioral reactions to intergroup interactions are naturally arising characteristics of normal human functioning (Allport, 1954; Dovidio et al., 2013; Ito, Thompson, & Cacioppo, 2004) and thus play a social function in groups (Macrae & Bodenhausen, 2000; Taylor, 1981), they can also impede social interactions, harm members of negatively-evaluated groups, and undermine equity in society. For these reasons, and acknowledging this caveat, as we present the research linking mindsets about malleability to stereotyping and prejudice, we largely consider these intergroup phenomena in terms of their negative or unwanted effects on individuals and society. As we will show, mindsets about malleability fundamentally shape each of these intergroup dynamics. For this reason, we advocate an approach to intergroup relations that considers people’s lay theories about malleability.

Mindsets About Malleability

This chapter focuses on mindsets about malleability, also called lay or implicit theories about malleability. Some people believe that characteristics (e.g., intelligence, personality, prejudice, groups, etc.) are fixed and stable—i.e., that they do not change over time. This belief is known as a fixed mindset, or entity theory. Other people, by contrast, believe that characteristics can grow and develop over time. This is called a growth mindset, or incremental theory. It is important to note that these beliefs, theories, or mindsets represent two ends of a continuous dimension along which people tend to be normally distributed (Dweck, 1999).

People’s lay theories or mindsets about the malleability of characteristics are naïve beliefs, learned through exposure and experience (Dweck, 1999). Because these beliefs focus on a core component of social understanding (i.e., malleability; Hong, Chiu, Dweck, Lin, & Wan, 1999; Weiner, Heckhausen, & Meyer, 1972; Weiner, 1985), they ground people’s meaning systems and snap into action as people encounter and interpret situations and other people (Plaks, Levy, & Dweck, 2009). In this way, lay theories serve as a perceptual lens or filter between a person and the world. That is, these beliefs drive people’s understandings of social information, their responses to it, as well as their goals and actions (Levy, Plaks, Hong, Chiu, & Dweck, 2001).

Mindsets are also domain specific (Dweck, Chiu, & Hong, 1995; Levy et al., 2001), pertaining to beliefs about the malleability of specific characteristics. The measurement of fixed versus growth mindsets involves asking people whether they agree or disagree with statements regarding a specific characteristic, such as

intelligence (e.g., “You have a certain amount of intelligence and you really can’t do much to change it”; Dweck et al., 1995; Dweck & Leggett, 1988), morality (e.g., “Whether a person is responsible and sincere or not is deeply ingrained in their personality. It cannot be changed much.”; Dweck et al., 1995), or prejudice (e.g., “People’s level of prejudice is something very basic about them that they can’t change very much”; Carr, Dweck, & Pauker, 2012), or regarding more generalized beliefs about the whole person (e.g., “kind of person” theories, “Everyone is a certain kind of person, and there is not much that they can do to really change that”; Chiu, Hong, & Dweck, 1997; personality theories, “A person can do things to get people to like them, but they can’t change their real personality”; Dweck, 1999), the nature of groups (“Every group or nation has basic moral values and beliefs that can’t be changed significantly”; Halperin, Russell, Trzesniewski, Gross, & Dweck, 2011; also see Rydell, Hugenberg, Ray, & Mackie, 2007), or the nature of the world (“Some societal trends may dominate for a while, but the fundamental nature of our world is something that cannot be changed much”; Dweck et al., 1995). Because mindsets are domain specific, it is possible for an individual’s beliefs to vary between strong incremental theories in some domains and strong entity views in other domains. In the review of research that follows, we will highlight the type of mindset that influences each intergroup phenomenon. We note, however, that the research on mindsets about malleability and intergroup relations has not been exhaustive, and therefore the question of which mindsets are most meaningful in intergroup dynamics and why remains open.

As naïve beliefs, implicit theories are latent knowledge structures that can be activated when relevant to the task or situation (Dweck & Leggett, 1988; Dweck, 1999). Here, the term “implicit” is used to evoke the fact that these beliefs underlie social perception and interaction, whether consciously or nonconsciously (Dweck, 1999)—as opposed to “implicit bias” which refers specifically to unconscious stereotypic associations (Greenwald & Banaji, 1995). Therefore, while people may not spontaneously describe their fundamental beliefs about the malleability or fixedness of a given characteristic as orienting their social understanding, they are nevertheless readily able to express these core beliefs when asked. Given this, measures of implicit theories are relatively direct and straightforward, as can be seen from the example items in the paragraph above.

We emphasize that these terms do not represent the sorting of people into different personality types. While most people have a core belief within a domain that drives their understandings, evaluations, and reactions, this does not mean that they are unaware of the alternative. Because of this general knowledge, it is possible to temporarily activate a specific theory in an experimental manipulation. Researchers have done so using articles that purport to summarize scientific findings (Bergen, 1991; Chiu et al., 1997), or with the subtler biased questionnaire manipulation that exposes participants to a target theory and only offers the option of agreement (Job, Dweck, & Walton, 2010; Rattan, Savani, Naidu, & Dweck, 2012). Research on mindsets about intelligence also shows that people can be trained in and convinced of a growth mindset with consequences that unfold over time, suggesting that such interventions can have lasting impacts on individuals’

beliefs (Aronson, Fried, & Good, 2002; Blackwell, Trzesniewski, & Dweck, 2007; Good, Aronson, & Inzlicht, 2003).

Most importantly, whether they are measured or manipulated, mindsets have meaningful consequences for stereotyping, prejudice, and people's reactions to intergroup contact. In the next section, we review research that showcases how mindsets (about the malleability of personality, of the "kind of person" someone is, and of prejudice) affect stereotyping and prejudice from the perspective of perceivers, that is, those who observe others across group boundaries and may exhibit stereotyping, prejudice, and discrimination. Following that, we summarize how targets of prejudice, that is, those who are subject to bias, are shaped by mindsets (about intelligence, personality, and groups) in contexts where they face stereotyping and overt or subtle prejudice. Of course, across situations the roles of perceiver and target can vary, and in some situations of intergroup conflict, people are both perceivers and targets at the same time (Richeson & Shelton, 2007; Shelton, Richeson, & Vorauer, 2006). Acknowledging this complexity, we use these terms to highlight whether we are discussing the expression of intergroup bias (i.e., on the part of perceivers), or responses to intergroup bias (i.e., on the part of targets). Toward the end of this chapter, we will return to discuss the broader implications of this work for intergroup reconciliation, which necessarily spans the perceiver-target distinction, and the potential for mindsets to promote more positive intergroup relations.

We also note that there are other types of lay theories that relate to intergroup dynamics (for a review, see Levy, Chiu, & Hong, 2006), but given our focus on mindsets about malleability, these are outside the scope of this chapter. To offer just a few examples of lay beliefs particularly relevant to the intergroup domain, research has documented the importance of lay theories regarding the protestant work ethic (Levy, West, Ramirez, & Karafantis, 2006; Rosenthal, Levy, & Moyer, 2011), beliefs about diversity (Plaut, Thomas, & Goren, 2009; Rattan & Ambady, 2013; Richeson & Nussbaum, 2004; Rosenthal & Levy, 2012), and belief in a just world (Bal & Van den Bos, 2017; Lerner, 1980). Others have also explored the related but distinct construct of psychological essentialism, defined as the belief that groups are distinct from one another because of their immutable, inherent and biology-based essences (Bastian & Haslam, 2006; Haslam, Rothschild, & Ernst, 2000; Haslam, 2017; Pauker, Ambady, & Apfelbaum, 2010), and lay theories about the biological or genetic nature of characteristics (Keller, 2005; Sanchez, Young, & Pauker, 2015; Williams & Eberhardt, 2008; Yzerbyt, Rocher, & Schadron, 1997). We note that holding a fixed mindset does not necessarily imply believing that characteristics are genetically determined, or denying the role of the environment in influencing these characteristics. It is indeed possible to believe that traits are influenced by one's environment and experiences over a certain period of time, but that after this period, traits come to "consolidate" into a fixed and stable state (Dweck, 1999). Similarly, holding a growth mindset does not necessitate rejecting the role that genetics may play in influencing individual characteristics, or equate to claiming that anybody has the potential to reach any goal (Dweck, 1999; Rattan, Savani, Naidu, & Dweck, 2012). More research should explore how these

different lay beliefs relate to one another, insofar as they do, and how they might mutually constitute stereotyping and prejudice (Haslam, Bastian, Bain, & Kashima, 2006; Levy, et al., 2006a, b).

Perceivers' Mindsets, Stereotyping, and Prejudice

How do intergroup dynamics unfold on the perceivers' side? It begins with categorization (Taylor, 1981), which happens rapidly and often automatically (Ito et al., 2004; Ito & Urland, 2003). As perceivers categorize a person into an outgroup, the content knowledge associated with the corresponding social category comes online (Crisp & Hewstone, 2007; Freeman & Ambady, 2009, 2011; Macrae, Bodenhausen, & Milne, 1995). This content knowledge represents cognitive associations formed about the group through the course of everyday interactions with the world, and is known as stereotype content knowledge (Fiske, Cuddy, Glick, & Xu, 2002). Whether endorsed or not, these social group associations can rise to the level of consciousness, or remain nonconscious, but in either case can influence attitudes, cognition, and behavior (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Greenwald & Banaji, 1995). When perceivers agree with negative stereotypes, they can also exhibit the affective negativity that is referred to as prejudice (outgroup hatred, as opposed to ingroup love, Brewer, 1999; de Dreu, 2010; Halevy, Bornstein, & Sagiv, 2008; Halevy, Weisel, & Bornstein, 2011). When negative stereotypic associations and affective negativity yield negative or biased behavior or policies toward outgroups, this is considered discrimination (Dovidio et al., 2013).

These are basic processes fundamental to intergroup perception. But we also know that there is variance in the degree to which people endorse stereotypes and exhibit prejudice (Dunton & Fazio, 1997; Fiske & Neuberg, 1990; Kawakami, Dion, & Dovidio, 1998; Livingston & Drwecki, 2007), and variance in how much these intergroup dynamics are cued by environments (Crandall, Eshleman, & O'Brien, 2002; Pettigrew, 1959). We suggest that an approach to intergroup relations that accounts for the influence of lay theories about malleability can offer insight into this variance. Although much is left to do, we highlight research that offers compelling evidence that beliefs about malleability affect categorization, stereotype formation and endorsement, as well as the expression of prejudice.

Categorization

Upon encountering a novel individual, people diagnose their social category group memberships (Ito et al., 2004; Taylor, 1981). In the context of person perception, a fixed mindset is associated with viewing even minimal information as indicative of a person's character (Chiu et al., 1997). Extending this perceptual tendency to the

intergroup context, people who believe strongly that the “kind of person” someone is cannot be changed (a fixed mindset, or entity theory) may consider social category membership as indicative of identity to a greater degree than people who believe strongly that the “kind of person” someone is can be changed (a growth mindset, or incremental theory). Consistent with this prediction, Eberhardt, Dasgupta, & Banaszynski (2003) documented a difference in how entity versus incremental theorists respond to racial categories. Participants received demographic information about a target person that listed this person’s race as either “Black” or “White.” On the next screen, the image of the target person presented was a computer-generated face, morphed from a Black original face and a White original face. The question was whether people with fixed versus growth mindsets would differ in how they applied the prior categorization. To assess this, later in the study, participants had to recall the target person that they had seen by choosing him from two images. Unbeknownst to participants, neither image actually represented the target person that they had seen. Rather, one had been morphed with more of the original Black face, and the other had been morphed with more of the original White face. In a first study, which measured participants’ lay theories, entity theorists were more likely to choose the category-congruent face, whereas incremental theorists were more likely to choose the category-incongruent face. That is, those with a fixed mindset assimilated their mental representation of the target person toward the category label initially mentioned in the demographic information, choosing the “more Black” face when the target person had been labeled Black and choosing the “more White” face when the (same) target person had been labeled White. In contrast, those with a growth mindset exhibited the opposite pattern, contrasting away from the category label by choosing the image more dissimilar to the category initially mentioned.

A second study that manipulated mindsets replicated these patterns, showing a causal relationship between mindsets about malleability and these consequences for social categorization. Moreover, participants were asked to draw, from memory, the person they had seen. Independent judges (who never saw the original face or the category label) rated the drawings made by those in the entity theory condition as more in line with the category label participants had seen, but rated the drawings made by those in the incremental theory condition as more in line with the category label *opposite* to the one that was seen. The fact that participants exhibited these patterns even though they were equally able to correctly recall the racial label suggests that this was not mistaken identity or misunderstanding. Rather, these findings suggest that when it comes to social group categorization, people across mindsets are equally capable of accurately categorizing, but those with a fixed mindset ascribe and adhere to categories to a greater degree, whereas those with a growth mindset focus more on the characteristics that deviate from the category.

Stereotype Formation

These consequences for social categorization raise the question of whether mindsets about malleability also shape the way in which people develop impressions of groups. Building upon previous lay theories research in person perception, which showed that a fixed mindset orients perceivers more toward diagnosing a person's character and maintaining this judgment (Chiu et al., 1997; Erdley & Dweck, 1993), Sheri Levy and her colleagues theorized that a fixed, rather than a growth, mindset about the kind of person someone is might predispose perceivers toward treating even minimal information as characteristic of the whole group. That is, these researchers tested whether mindsets about malleability shape people's predispositions to forming stereotypes about groups. They offer compelling, convergent evidence from both middle school children (Levy & Dweck, 1999) and adults (Levy, Stroessner, & Dweck, 1998) showing that mindsets shape stereotype formation.

To test this proposition, the researchers offered an opportunity for perceivers to form a negative stereotype about another group. Children (aged 11–13) read about another (fictitious) school, in which several different students behaved in mostly negative ways such as calling a classmate's artwork ugly or not helping a classmate who dropped papers (Levy & Dweck, 1999). The question was to what degree children who believed personality is fixed versus malleable would stereotype this school as negative, and whether they would apply the stereotype to all students at the school. As theorized, the children who held a fixed mindset about personality formed more extreme negative stereotypes (on average considering the students at the school to be "mean") compared to the children who held a growth mindset about personality (on average considering the students at the school to be "a little mean"). Entity theorist children also applied these characterizations more globally, i.e., to all students in the school, than did incremental theorist children (Levy & Dweck, 1999). The difference in the stereotype formation process was perhaps most compellingly captured in children's verbal explanations for why students at the school behaved as they did. Entity theorists offered explanations grounded in traits more often, whereas incremental theorists offered explanations that touched on external factors. An additional study revealed that entity and incremental theorists differ in the degree to which they develop both positive and negative stereotypes. When asked to rate the students at the schools on a series of traits (nice–mean, honest–dishonest, friendly–unfriendly, generous–stingy, good–bad), fixed mindset children exhibited more extreme stereotypic judgments for both the "good" and the "bad" school compared to growth mindset children. That is, fixed mindset children perceived students at the "good" school significantly more positively and students at the "bad" school significantly more negatively than growth mindset children. In addition, fixed mindset children perceived students from each of the two schools as less likely to share interests and activities with each other than did growth mindset children (Levy & Dweck, 1999). The fact that entity theorists also formed more positive stereotypes for the "good" school compared to incremental theorists

suggests that they are not necessarily more negative or critical than incremental theorists (and vice versa, that incremental theorists are not just optimists compared to entity theorists), but rather that a fixed mindset predisposes people to form and rely on stereotypes more than a growth mindset.

Adults also show this predisposition toward stereotype formation when they hold more fixed views of others (Levy et al., 1998). Undergraduates read about student groups at another school, who (among some neutral behaviors) either engaged in negative behaviors, such as pushing to the front of a line, or in positive behaviors, such as sharing an umbrella with a stranger. They then wrote descriptions of the groups. Entity theorists made more references to traits in their descriptions of the groups, used more extreme adverbs (e.g., “very,” “always”), and reported seeing the “good” and “bad” groups as more dissimilar compared to incremental theorists. Particularly noteworthy is the fact that entity theorists made their ratings of the groups faster than incremental theorists, and reported feeling more satisfied with the (relatively minimal) amount of information they had been offered about these groups. That is, not only were entity theorists more likely to spontaneously form extreme stereotypes about the groups, but they were also quicker to form their judgments compared to incremental theorists.

Research also shows that mindsets play a role in protecting stereotypes. That is, extensive research has shown that stereotypes, once formed, are highly resistant to change (Bodenhausen, 1988; Devine & Elliot, 1995; Maass, Salvi, Arcuri, & Semin, 1989; Macrae, Hewstone, & Griffiths, 1993; Weber & Crocker, 1983). Plaks, Stroessner, Dweck, & Sherman (2001) explored what role mindsets about malleability might play in the maintenance of stereotype content. The researchers both measured and manipulated people’s mindsets about the malleability of the kind of person someone is. They exposed people to stereotype-consistent information, stereotype-inconsistent information, or stereotype-irrelevant information about a Nazi (a negatively-stereotyped target) or a priest (a positively-stereotyped target), or, in another study, an entirely novel social group. In line with the findings reviewed above, those who held fixed mindsets clung to stereotype-consistent information, paying it more attention and preferentially remembering it. Those with growth mindsets, by contrast, allocated more attention to and remembered information that differed from the stereotype, whether it was inconsistent or irrelevant. Thus, a picture begins to form of how the same world may appear strikingly different from the entity versus incremental perspective (Dweck et al., 1995). Those who believe personality and people are fixed are more likely to attend and adhere to category boundaries, form more extreme stereotypes more quickly, and preferentially attend to and remember stereotype-consistent information relative to those who believe personality and people are malleable.

Stereotype Endorsement

Even if those who hold a fixed mindset are more likely to form associations between groups and characteristics, it does not mean that they necessarily endorse social stereotypes to a greater degree; indeed, one can know the content of a stereotype while disagreeing with it (Devine, 1989). Levy, Stroessner, & Dweck (1998) tested this directly. When they asked people with growth and fixed mindsets to list cultural stereotypes that exist about different groups in society, including African Americans, Asians, Caucasians, Hispanics-Latinos, and Jews, there were no differences in the number or valence of stereotypes listed. That is, when it comes to stereotype knowledge, people across the range of mindsets are equally likely to know stereotypic associations that are prevalent in society. Yet, when they tested for stereotype agreement – how much these participants considered the stereotypes to be true – systematic differences emerged. Those who had endorsed a growth mindset were significantly less likely to agree that the stereotypes were true compared to those who had endorsed a fixed mindset (Levy et al., 1998).

Another study confirmed that mindsets have a causal impact on stereotyping (Levy et al., 1998). Participants were randomly assigned to read an article that described scientific evidence either stating that “personality is changeable and can be developed,” (the incremental theory condition) or that “personality, like plaster, is pretty stable over time” (the entity theory condition; Chiu et al., 1997). These articles manipulated participants’ beliefs about the malleability of personality, at least for the short term. Participants then indicated how much they thought different characteristics accurately described African Americans, Asians, Latinos, teachers, doctors, lawyers, and politicians. Some of the characteristics represented stereotypes of the groups, and some did not. Participants who had been randomly assigned to read the entity theory article were more likely to agree that the relevant stereotypic characteristics were descriptive of these racial and occupational groups than those who had read the incremental theory article. There was no difference in how descriptive participants rated the non-stereotypic characteristics, showing that mindsets have particular relevance to how people apply social stereotypes, not social descriptors in general. While stereotyping is multiply determined, Levy et al. (1998) have shown that the influence of mindsets about the malleability of personal characteristics on stereotyping occurs over and above the influence of social desirability, right-wing authoritarianism, attributional complexity, need to evaluate others, and personal need for structure. In sum, this body of work reveals that entity and incremental theorists endorse societal stereotypes to different degrees.

Prejudice

If people’s malleability mindsets shape how much they assimilate individuals into social categories, how readily they form stereotypes, and how much they endorse

stereotypes, then do they also drive affective negativity toward stigmatized groups (i.e., prejudice) or differential behavior toward stigmatized group members (i.e., discrimination)? Classic research on mindsets and intergroup relations has included hallmark measures of prejudice, such as a (lack of) affective warmth (e.g., feeling thermometer; Krysan, 2000; McConahay, 1986), desire to maintain social distance (Bogardus, 1947), and insensitivity toward the suffering of disadvantaged groups (Čehajić, Brown, & González, 2009; Harris & Fiske, 2006). This research offers suggestions of a role for mindsets about malleability in the expression of intergroup prejudice.

Recall that Levy et al. (1998), asked undergraduates to read about student groups at another school and varied whether the groups were described as engaging in positive or negative behaviors. In addition to the findings reported above, the researchers included a measure similar to classic prejudice measures of affective negativity (McConahay, 1986). They found that those with a fixed mindset evaluated the negative group more negatively than those with an incremental mindset on a scale ranging from “very negative” (−100) to “very positive” (100; Levy et al., 1998), suggesting that a fixed mindset might predispose people to exhibiting greater affective negativity toward outgroups than a growth mindset.

Similarly, Levy & Dweck’s (1999) study of middle school children’s stereotype formation included a measure of willingness to interact with members of the novel groups that children learned about, reminiscent of classic measures of social distancing (Bogardus, 1947). While children generally were disinclined to interact much with a group who behaved in an undesirable manner, entity theorists were still more likely to socially distance compared to incremental theorists, reporting less desire to attend a party or be friends with members of the novel group (Levy & Dweck, 1999).

Outgroup prejudice is also known to impair helping (Cuddy, Rock, & Norton, 2007; Hornstein, 1978; Levine, Prosser, Evans, & Reicher, 2005). Karafantis and Levy (2004) therefore explored whether mindsets about malleability would play a role in children’s outlook toward helping homeless children. They found that 9–12-year-old children who believed human attributes were more malleable had more positive attitudes toward homeless children, were more open to social interactions with homeless or UNICEF-funded children, and reported more past volunteering for people in need compared to children who believed human attributes were more fixed. When given the volunteering opportunity of participating in the Trick-or-Treat for UNICEF Program, children with growth mindsets about human attributes reported being more active participants, enjoyed the experience more, and were more willing to help in the future than children with fixed mindsets (Karafantis & Levy, 2004). Given the correlational nature of these results, it is of course difficult to determine the direction of these effects; it could be that volunteering engendered a more growth-oriented perspective among children, which in turn promoted active participation, enjoyment of volunteering, and willingness to offer future help. While Karafantis & Levy’s (2004) results suggest that mindsets could play a role in sustaining a virtuous circle of diminishing prejudice toward

outgroups through social engagement with causes, more research ought to be conducted to test the causal pathways.

The studies reviewed above are indicative of the idea that mindsets about malleability play a role in the expression of intergroup prejudice, although more must be done to investigate this directly. More recently, research has also investigated the role of fixed and growth mindsets in explaining the occurrence of discrimination, focusing on the persistent issue of gender discrimination in the workplace (Rudman & Kilianski, 2000; Simon & Hoyt, 2008). Consistent with previous work in the domain of stereotyping and prejudice (Levy et al., 1998, Levy & Dweck, 1999), Hoyt and Burnette (2013) found that participants with a growth mindset were less likely to exhibit a stereotypic preference for male versus female authorities. Going further, they found that perceiving the agentic leader prototype as more congruent with males (rather than females) is less predictive of discriminatory evaluations of female leaders among incremental theorists than it is among entity theorists (Hoyt & Burnette, 2013). That is, although everyone on average endorsed the prevalent and persistent gender stereotypes about leadership being a purportedly masculine attribute to some degree, those stereotypical associations yielded differential treatment of female leaders more among employees who held fixed, rather than growth, mindsets (Hoyt & Burnette, 2013).

Although more direct investigations of the link between mindsets about malleability and prejudice are necessary, the existing evidence suggests that those who believe personal attributes to be fixed exhibit more negative attitudes toward stigmatized groups, report more desire to maintain social distance from them, and are less motivated to offer help or contribute to those groups' improvement. Similarly, additional research should explore how these mindsets influence different types of real-world discrimination, both in the domain of gender and management (Hoyt & Burnette, 2013) and beyond.

Lay Theories of Prejudice

The preceding section showcased research that linked mindsets about malleability to prejudice, both in terms of its classic definition of negative animus and through its indicators, particularly avoidance of social interactions. However, more recent research suggests a more nuanced perspective is essential. Depending upon one's mindset, the prosocial desire to avoid being prejudiced against outgroups can ironically engender the precise avoidance or awkwardness in social interactions that typically are indicators of prejudice. How can this be? Note that all of the research reviewed above focused on lay theories about whether personality or the "kind of person" someone is can change or not. Earlier, we highlighted that people can also have mindsets about the malleability of domain-specific characteristics. Specifically, research on people's beliefs about the malleability of *prejudice* itself has discovered these surprising and ironic consequences.

In the preceding sections, we highlighted that perceivers are those who observe and classify others, in the course of which stereotypic associations can be activated and prejudice can be expressed. What about members of majority groups who *do not* endorse, or do not want to exhibit, stereotyping and prejudice? Carr et al. (2012) and Neel and Shapiro (2012) point out that, for these individuals, intergroup situations can represent a performance context. Following classic work in the domain of beliefs about intelligence (Dweck & Leggett, 1988), these researchers theorized that perceivers who view prejudice as fixed might have a greater performance orientation. That is, those with fixed mindsets about prejudice might want to showcase their unbiased nature and avoid challenging situations that might call their beliefs into question, which might lead them to avoid or exit situations in which they might reveal bias. In contrast, they theorized that perceivers who view prejudice as malleable might have a more learning oriented perspective, leading them to approach intergroup situations with an open outlook and to engage with challenge in this domain. Across studies where they both measured and manipulated these lay theories of prejudice, the researchers found support for this theory. Indeed, participants who endorsed a more fixed view of racial prejudice wanted less information about bias, wanted to and did exit intergroup interactions more readily, and were less likely to take learning opportunities focused on race and racism (Carr et al., 2012; Neel & Shapiro, 2012). When Carr et al. (2012) asked fixed and growth mindset participants to set up chairs for a conversation with an outgroup member, fixed mindset participants set the chairs almost 10 inches further away from each other than growth mindset participants did. These beliefs also accounted for awkwardness *during* interracial interactions; perceivers with fixed mindsets exhibited more anxious, negative, and disengaged nonverbal behaviors in an in-person interaction that either focused on race or involved an outgroup member (Carr et al., 2012; Neel & Shapiro, 2012). These effects emerged above and beyond participants' degree of racial bias, highlighting that similar types of awkwardness and social disengagement in intergroup interactions might arise from negative animus toward outgroups, as well as, paradoxically, from a belief that prejudice is fixed and the performance-oriented outlook that follows (Carr et al., 2012; Neel & Shapiro, 2012).

This research on lay theories about prejudice converges with other work in intergroup relations highlighting that a focus on performance in interracial interactions can heighten anxiety and tension, and reduce the fluency of such encounters (Butz & Plant, 2009; Richeson & Trawalter, 2005). Classic research suggests that learning orientations stem from growth mindsets, while performance orientations follow from fixed mindsets (Dweck & Leggett, 1988). Migacheva and Tropp (2012) offer compelling evidence of this link. They assessed how much middle and high school students thought they could learn from members of outgroups (i.e., an intergroup learning orientation; Migacheva & Tropp, 2012). They found that for both European and African American middle school children, having a learning (versus performance) orientation predicts greater levels of comfort and interest in engaging in interactions with members of the other group (Migacheva & Tropp, 2012). These beliefs affected interactions over time as well. Students who expressed

a learning orientation three weeks before a community-focused diverse summer camp expressed more comfort with, and interest in, interacting with members of different groups at the end of the camp (Migacheva & Tropp, 2012). Together with the work on mindsets about prejudice, these results suggest that promoting a learning orientation, or growth mindset, can be an effective lever to improve intergroup relations in the long run.

Summary

Mindsets about malleability influence multiple stages of intergroup processes on the perceivers' side. Mindsets about personality and the "kind of person" someone is shape categorization, stereotype formation, endorsement and maintenance, as well as the expression of prejudice. Yet, mindsets about prejudice itself offer new insight into why prejudiced-seeming behavior might emerge despite positive intentions on the perceiver's side. This body of evidence highlights that a full understanding of intergroup relations on the perceiver's side necessitates a consideration of mindsets about malleability.

Targets' Mindsets and Responses to Stereotyping and Prejudice

Turning from the expression of stereotyping, prejudice, and discrimination to the experience of it, we now focus on mindsets about malleability and their consequences among targets of bias. Research has extensively documented the negative consequences of both subtle and overt stereotyping and prejudice for members of stigmatized groups (Dardenne, Dumont, & Bollier, 2007; Vescio, Gervais, Snyder, & Hoover, 2005; Wolfe & Spencer, 1996). Although there is much left to explore, research to date suggests that mindsets about malleability can play a role here too, influencing targets' vulnerability to stereotypes and their reactions to overt expressions of prejudice.

Social Identity Threat

A concern about confirming negative stereotypes about one's group, known as stereotype threat or social identity threat, characterizes the experiences of stigmatized individuals in diagnostic performance contexts where stereotypes are salient (Steele & Aronson, 1995). This sense of threat has a myriad of negative consequences: increased anxiety, reduced working memory and learning, reliance on the

first response that comes to mind during a test, and increased rumination (Jamieson & Harkins, 2007; Schmader, Johns, & Forbes, 2008; Schmader & Johns, 2003; Taylor & Walton, 2011). More broadly, stigmatized individuals under social identity threat exhibit decrements to their performance and their sense of belonging (Good, Rattan, & Dweck, 2012; Steele & Aronson, 1995).

Considering the conditions under which stereotype threat arises, Aronson, Fried, and Good (2002) considered the possibility that a growth mindset about intelligence might serve an important buffering role. Recall that social identity threat arises when stereotypes are salient in diagnostic performance conditions. Yet, diagnostic situations are not evaluated or experienced in the same way by everyone. In her seminal work on mindsets, Dweck (1999) (see also Dweck & Leggett, 1988) showed that people who view intelligence as fixed orient toward performance and see difficulty as indicative of a lack of ability, whereas people who think intelligence can grow orient toward learning and see difficulty as an opportunity to overcome challenge through effort. Pairing these insights, Aronson et al. (2002) investigated whether stigmatized students armed with a growth mindset might exhibit less vulnerability to the performance decrements associated with stereotype threat.

In the malleable pen pal condition, Aronson et al. (2002) taught undergraduate students about the growth mindset about intelligence. Students first watched an instructional video on the latest scientific evidence suggesting that brain capacities can grow, and then were asked to write a letter conveying this message to an “at risk” middle school student, using the ideas they had been exposed to and examples from their personal life. Participants in the control pen pal condition watched a video clip describing scientific evidence pointing to the multi-faceted nature of intelligence, and wrote a letter to a pen pal conveying this message (Aronson et al., 2002). In a second session, all participants again wrote the target message (either describing intelligence as malleable or multi-faceted, depending upon their condition) to a second pen pal. In the third session, participants were asked to transform their messages into audiotaped speeches, to be used in future interventions in schools. This three-session format was constructed in order to offer a strong condition manipulation. There was also a true control “no pen pal” condition, in which participants did not participate to any of the activities of these three sessions (Aronson et al., 2002).

Nine weeks later, African American students in the growth mindset intervention condition reported more enjoyment of academics, more identification with academic achievement, and better academic performance (controlling for prior SAT scores), compared to those in the two control conditions. While on average all students in the malleable condition exhibited benefits from learning about the growth mindset, the benefits were most striking among African Americans, the group vulnerable to stereotype threat in the academic context. Extending this work, a field experiment showed the impact that a growth mindset intervention could have on improving standardized test scores among students stigmatized by stereotypes about their gender, race, and income (Good et al., 2003). The researchers induced a growth mindset about intelligence in low-income, largely Latino/Hispanic

seventh-graders through the guise of a computer skills course. Girls in the growth mindset condition (as well as in an attributional retraining condition, and a combined growth mindset + attributional retraining condition) performed significantly better on the math portion of their end-of-year state standardized testing than girls in an anti-drug control condition.

Good, Rattan, and Dweck (2012) found that mindsets about the malleability of math intelligence not only affect performance, but they also influence the sense of belonging to math among stigmatized students. Recognizing that broad-scale stereotypes about ability are persistent and pervasive in academic settings, the researchers explored undergraduate women's sense of belonging as it unfolded over the course of a semester-long college-level calculus class. Moreover, this research investigated the possibility that it is not only one's *own* mindset that matters, but also the mindsets one perceives among important others in the context – in the case of an academic context, the teachers and other students. A few weeks into this semester-long calculus course, Good et al. (2012) measured women's perceptions of stereotyping in the classroom environment (e.g., "People in my calculus class believe that females are as good as males in calculus"), as well as their perceptions of a fixed versus growth mindset about math ability in the classroom environment (e.g., "People in my calculus class believe that people have a certain amount of math intelligence and they can't really do much to change it"). They found that when women perceived higher levels of gender stereotyping in the classroom context, also perceiving a growth (rather than fixed) mindset in the classroom protected their sense of belonging to math. Indeed, a highly stereotyping environment paired with growth mindset messages left women's sense of belonging as high as the sense of belonging evidenced among women who reported being in relatively low stereotyping environments. This protection had important consequences, since their higher sense of belonging to the math domain engendered a greater desire to pursue math in the future and higher end-of-term math grades.

Extending this theory into the domain of employment, Emerson and Murphy (2015) investigated whether organizations that espouse a fixed (versus growth) mindset about ability are more threatening to stigmatized employees. Mindsets about intelligence were conveyed through a manipulation of corporate mission statements. In the entity condition, a consulting company described its "performance-oriented" mission to recruit candidates with the "best" instincts and ideas, and to help employees be "the geniuses they are" by "encouraging, recognizing, and rewarding intelligence." In the incremental condition, the consulting company's mission statement was described as a "growth-oriented" one that involved recruiting motivated candidates with "a love for learning, passion, creativity and resourcefulness," as well as to help employees "improve and push through limits" by "encouraging, recognizing, and rewarding development." Women who read the mission statement that communicated an organization's fixed view of intelligence reported less trust in the organization than women who read the growth-oriented mission statement (Emerson & Murphy, 2015). The authors offered direct evidence of the link between environments that communicate fixed views of intelligence and heightened threat: the mistrust engendered by entity-oriented

environments (compared to the incrementally-oriented environments) was driven by women's heightened concerns about being negatively stereotyped. These processes are theorized to emerge because of the heightened diagnosticity of situations in an entity worldview. Consistent with that, women who imagined performing poorly in a meeting with a company representative were more likely to disengage, but only when they were told the organization espoused fixed, rather than growth, views of ability.

These findings are supplemented by other research, which has found that when abilities are characterized as genetic or innate (expressions of a fixed view), gender differences in performance and pursuit ensue. Dar-Nimrod & Heine (2006) found that characterizing math ability as genetic (a view congruent with fixed mindsets) led women to underperform on a math test relative to women for whom math ability had been characterized as environmentally-determined (a more growth-congruent view). Leslie, Cimpian, Meyer, & Freeland (2015) surveyed faculty in the academic fields of science, technology, engineering, and mathematics (STEM) and found that the more faculty viewed talent in their field as stemming from innate factors (i.e., a field-specific fixed view of ability), the fewer women Ph.D. recipients there were in the field. Indeed, Cimpian and colleagues have documented that simply attaching high performance on a challenging task to a specific social category yields underperformance among other groups, and they theorize that this underperformance arises due to the spontaneous formation of fixed views of the ability that underlies the task (Cimpian, Mu, & Erickson, 2012). Interestingly, Mendoza-Denton, Kahn, & Chan (2008) showcase the other side of these dynamics that link mindsets about the malleability of intelligence and the performance of stereotyped groups. When one's group benefits from the stereotypes (e.g., Asians in the case of math), an entity theory about ability can systematically *boost* performance relative to an incremental theory (Mendoza-Denton et al., 2008).

In sum, being targeted by negative stereotypes has adverse consequences for stigmatized individuals' performance, sense of belonging, and overall engagement with the field in which they are stereotypically expected to do poorly. However, targets' vulnerability may depend upon the mindsets about intelligence or about specific domains of study (e.g., math) that they hold, or those that the environment communicates. Members of stigmatized groups who held, were taught, or perceived growth mindsets showed less vulnerability to social identity threat and retained their sense of belonging more. This body of research suggests incremental views of abilities may be an important, but underused, intervention strategy for allowing stigmatized individuals' talent to thrive in classrooms and workplaces.

Responses to Overt Prejudice

The experience of prejudice in the modern age is not only composed of subtle and systemic stereotypes. In their everyday social and workplace interactions, members of stigmatized groups continue to face overt expressions of prejudice (Deitch et al.,

2003; Dixon, Storen, & Van Horn, 2002; Swim, Hyers, Cohen, & Ferguson, 2001; Swim, Hyers, Cohen, Fitzgerald, & Bylsma, 2003). This occurs when women and minorities are told explicitly that, for example, they lack competence in specific fields, do not belong in certain contexts, or are unfit to lead due to their group memberships (Ely, Meyerson, & Davidson, 2006; Sue, 2010). Experiencing such overt bias has profound negative consequences for the psychological (Feagin & Sikes, 1994; Richeson & Shelton, 2007; Williams & Williams-Morris, 2000) and physiological (Harrell, Hall, & Taliaferro, 2003; Sawyer, Major, Casad, Townsend, & Mendes, 2012; Williams, Neighbors, & Jackson, 2003) outcomes of members of targeted groups. When faced with overt prejudice, women and minorities report wanting to speak out to express their disagreement with it, but often are held back from doing so due to situational pressures (Swim & Hyers, 1999) or the real risk of social and professional costs (Woodzicka & LaFrance, 2001). When they remain silent in the face of bias, women and minorities experience negative self-directed emotions and regret (Shelton, Richeson, Salvatore, & Hill, 2006).

Rattan and Dweck (2010) considered whether the mindsets that targets of prejudice hold might shape their experiences with overt prejudice. They theorized that members of stigmatized groups who hold a fixed mindset might view someone who expresses overt prejudice as fundamentally bad (e.g., racist or sexist). Because a fixed view of others means that someone who is biased will remain biased, this perspective might make responding to an expression of overt prejudice seem unlikely to have an impact. In contrast, they theorized that members of stigmatized groups who believe others can grow and develop might be less likely to diagnose someone as fundamentally biased based on a single expression of prejudice. Then, to those with growth views, the confrontation of prejudice might even represent an opportunity to educate and ameliorate perpetrators of bias. The researchers focused on implicit theories of personality, specifically, women's and minorities' views of whether personality can change. When minority undergraduates interacted with a confederate who expressed overt bias, those who held growth views of others' personality were more likely to spontaneously speak out to express their disagreement with the biased statement than those who held fixed views (Rattan & Dweck, 2010). In other studies, which assessed minorities' and women's responses to scenarios in which they encountered bias, the researchers found that participants with fixed versus growth mindsets (whether measured or manipulated) reported being less willing to confront the statement, even though they disagreed equally, and being less willing to interact with the perpetrator of bias.

While targets of prejudice should never bear the burden of being expected to address prejudice, the extant research suggests that they often want to speak out to express their perspectives (Shelton et al., 2006a, b; Swim & Hyers, 1999). From this perspective, we can see that a growth mindset affords minorities and women a motivational basis to take this desired action and to keep an open mind toward future relations subsequently (Rattan & Dweck, 2010). But what happens if growth mindset targets of prejudice do not speak out in the face of bias, which is not just possible but expected given the many factors that can silence women and minorities who face overt bias (Swim & Hyers, 1999; Woodzicka & LaFrance, 2001)? Earlier,

we pointed out that growth versus fixed mindsets do not simply map on to optimism versus pessimism, respectively. Instead, mindsets are a complex lens through which situations and actions are interpreted. Given this, absent a concrete action that offers perpetrators of bias an indication of (and an opportunity for) change, will growth mindsets still yield a positive outlook following an encounter with explicit bias?

In more recent work, Rattan and Dweck (2016) have tested exactly this question. They found that when minorities and women who held growth mindsets anticipated staying silent in the face of prejudice, they had an equally *negative* outlook on the perpetrator of bias as did minorities and women who held fixed mindsets. However, when fixed and growth mindset participants equally anticipated confronting prejudice, only those with a growth mindset experienced the benefits of a more positive outlook on the perpetrator of bias and maintained their sense of belonging and workplace satisfaction to a greater degree. Those who spoke out but had a fixed mindset were as negative in their outlook toward the perpetrator of bias, and reported similarly low sense of belonging and workplace satisfaction as those who stayed silent. These findings suggest that a growth mindset may offer adaptive advantages to minorities and women who face overt expressions of prejudice, but only when these lay theories are paired with change-oriented behaviors, such as the confrontation of prejudice.

Intergroup Reconciliation

Rattan & Dweck's (2010, 2016) results suggest that growth mindsets may be essential to addressing and reconciling after an instance of daily overt stereotyping or prejudice. In contrast to the everyday bias discussed above, protracted (or intractable) conflicts correspond to conflicts that involve a long history of rivalry and failed attempts at peace-making (Bar-Tal, 2001; Coleman, 2003; Vallacher, Coleman, Nowak, & Bui-Wrzosinska, 2010). In such hostile contexts, stereotype endorsement and expression are commonplace (Bar-Tal & Halperin, 2011), and the roles of perceiver and target fluctuate owing to the reciprocally negative attitudes that both sides hold toward each other (Halperin et al., 2011). Researchers have investigated whether mindsets can play a role in de-escalating such intractable conflicts (also see work on neutralizing interpersonal conflicts, Yeager, Miu, Powers, & Dweck, 2013; Yeager, Trzesniewski, Tirri, Nokelainen, & Dweck, 2011).

In the context of protracted conflicts, the natural tendency to interpret others' behaviors in terms of their dispositions, rather than of the situational pressures that they experience (a tendency called the "fundamental attribution error"; Ross, 1977) represents a particularly meaningful barrier to reconciliation. These dispositional attributions indeed communicate that the root of all evil deeds is in the other side's very nature, which can lead to an escalation of the conflict. However, Levontin, Halperin, and Dweck (2013) proposed that people's long-term attitudes toward the outgroup may be differentially affected by people's short-term attributions (whether

dispositional or situational), depending of their mindsets. They theorized that people who hold an incremental view of personality may be less influenced by short-term attributions when forming long-term attitudes toward the outgroup, because they believe that outgroup members' personality, just like their circumstances, can change. In contrast, people holding an entity theory of personality may be particularly likely to translate short-term attributions into long-term attitudes. Since entity theorists believe that personality, unlike circumstances, cannot change, believing that the outgroup's behavior stems from its very nature will have more serious implications than believing that it is the product of specific circumstances. The authors therefore expected that entity theorists' long-term attitudes toward the outgroup would be more affected by these short-term attributions to internal dispositions than those of incremental theorists.

Among Jewish Israeli participants, Levontin, Halperin, and Dweck (2013) independently manipulated lay theories of personality as fixed or malleable and the type of attributions, as dispositional or situational, that participants had to make to interpret behaviors of seven fictitious characters. Results showed that when participants were led to believe in an entity theory of personality and led to make dispositional attributions, they subsequently exhibited more negative stereotyping and less support for the civil rights of Israeli Arabs compared to their counterparts who had been led to adopt an entity theory and to make situational attributions. They also opposed compromising to resolve the longstanding conflict more. However, the relative differences in stereotyping, support for civil rights, and willingness to compromise across attribution conditions were essentially erased when participants were instead led to believe in an incremental theory of personality, suggesting that incremental theorists are indeed less influenced by the type of short-term attributions for outgroup behaviors than entity theorists. These results suggest that promoting the view that human nature is malleable can be one lever to promote more constructive intergroup relations.

Another way to improve protracted intergroup conflict may be to facilitate intergroup contact (Allport, 1954; Pettigrew & Tropp, 2006; Tropp & Pettigrew, 2005a, b). However, intractable conflicts typically offer environments that are not conducive to cross-group interactions (Crisp, Husnu, Meleady, Stathi, & Turner, 2010). In the context of the protracted conflict between Greek and Turkish Cypriots, Halperin et al. (2012) investigated the possibility that mindsets about group malleability might play a causal role in increasing willingness to engage in intergroup contact. Turkish Cypriot participants were assigned to read an article that described the negative behaviors of groups engaged in violent conflicts as either fixed or changeable over time. Afterwards, Turkish Cypriots who learned that groups are malleable reported significantly greater willingness to interact with a Greek Cypriot compared to their counterparts in the fixed condition. This was due to the lower anxiety experienced by those in the incremental theory (versus entity theory) condition. Particularly noteworthy is the fact that the mindset manipulation was effective in the absence of any specific mention of the Cypriot conflict. Therefore, these results raise the possibility that mindset interventions describing the malleable

nature of intergroup conflict *in general* can lower group-level anxiety and promote intergroup contact among parties involved in *particular* conflicts.

To overcome longstanding conflict and pursue peace-making processes, people need hope (Bar-Tal, 2001; Moeschberger, Dixon, Niens, & Cairns, 2005). What role might mindsets about malleability play in engendering hope among members of groups involved in intractable conflicts? Saguy and Halperin (2014) found that Israelis who saw a Palestinian expressing intra-group critiques felt hope and openness, but only when the Israelis held growth mindsets about groups. Indeed, Israelis who saw groups as fixed remained unaffected.

Cohen-Chen, Halperin, Crisp, and Gross (2014) measured and manipulated beliefs in the malleability of *conflicts* in general, and found that Jewish Israeli participants who saw (or were led to see) conflicts as more changeable reported significantly greater hope regarding the resolution of the Israeli–Palestinian conflict. Their greater sense of hope in turn led them to report significantly greater willingness to make concessions on the core issues of the conflict, relative to Jewish Israelis who viewed conflicts as fixed and unchanging. Cohen-Chen, Crisp, and Halperin (2015) found strikingly similar consequences for the experience of hope and willingness to compromise when investigating people’s beliefs about the malleability of the world.

In sum, mindsets about malleability can play a role in fostering positive group emotions (hope) and alleviating negative ones (anxiety), thereby creating the conditions necessary for positive intergroup outcomes to emerge. However, positive progress may also depend on the willingness to acknowledge and apologize for past wrongs, or to offer collective apologies (Lazare, 2004). Wohl et al. (2015) proposed that lay beliefs about the malleability of groups may play a role in explaining when people are more versus less open to collective apologies. Israelis with growth mindsets about groups reported being more ready to accept a collective apology from Palestinians for the killing of innocent Israeli civilians, and were more in favor of initiating a peace process than Israelis with fixed mindsets about groups. This was because a growth mindset cast the apology as indicative of remorse to a greater degree than did a fixed mindset. Importantly, the authors ruled out the possibility that people who report having a growth mindset may also simply be more forgiving in general in the absence of a collective apology, there was no difference in fixed and growth mindset participants’ willingness to forgive. Finally, as noted, transgressions are committed on both sides in protracted conflicts. Wohl et al. (2015) also found participants with a growth mindset significantly more willing to reciprocate the apology compared to participants with a fixed mindset.

In sum, even in the context of the most entrenched real-world group conflicts, lay theories about malleability (of groups, conflicts, or the world) have a role to play. Convincing each party to adopt an incremental view may represent an effective lever to change the way both sides look at each other, attenuate negative group-based emotions, enhance positive ones, and move toward more constructive peace processes.

Conclusion: A Mindset Approach to Intergroup Relations

Mindsets about malleability are meaning systems that function as an interpretive lens. For this reason, people's growth and fixed mindsets have the potential to influence virtually every aspect of the psychology underlying intergroup dynamics. On the perceivers' side, evidence shows that mindsets can shape categorization, stereotype formation, maintenance and endorsement, as well as the expression of prejudiced behavior (whether driven by negative animus or a desire to avoid prejudiced behavior). For those who are subject to negative stereotypes, mindsets can shape their vulnerability to social identity threat and responses to overt bias. Mindsets also have a role to play in longstanding real-world conflicts, offering insights into how to pave the way toward productive peace processes.

The study of intergroup relations gains much from considering perceivers' and targets' mindsets about malleability. While not yet exhaustive, the body of evidence on mindsets and intergroup relations challenges core assumptions in the study of intergroup dynamics. Consider the assumption that stereotyping is an inevitable, natural cognitive process. In the context of research on mindsets about malleability, we see that not all stereotyping is inevitable. To the contrary, growth mindsets orient perceivers toward individuating and category-inconsistent information (Eberhardt et al., 2003; Levy et al., 1998). Reflect on the behavioral indicators of prejudice according to the field of intergroup relations, such as the withdrawal from intergroup contexts and distancing from outgroup members. People's mindsets about prejudice reveal that these very same behaviors can actually arise out of a desire to avoid exhibiting bias, when people take a fixed perspective. More generally, research on intergroup dynamics has focused on biased cognition, attitudes, and behavioral reactions in the context of specific social groups. The research on mindsets about malleability highlights that generalized lay theories about personality, kinds of people and intelligence, which on the surface may seem to have no relevance to intergroup contexts, can yet have profound consequences for stereotyping and prejudice expression on the part of perceivers, and for targets' responses to both subtle and overt bias.

The field of intergroup relations has long documented the difficulties of creating meaningful change in people's stereotyping and prejudice, and the even greater challenge of maintaining it over time. Given this, another major contribution that the study of mindsets about malleability offers to the field of intergroup relations is the potential to implement concrete, practical, and scalable interventions that can be used to reduce group-based disparities, such as racial and gender achievement gaps, or to resolve longstanding conflicts. Though more research is of course necessary, particularly research that examines the long-term consequences of mindset interventions, integrating the study of lay theories into issues of intergroup relations offers untold potential for real-world impact.

Conversely, the study of mindsets gains much from considering intergroup relations contexts. The study of mindsets about malleability began with the investigation of the role of lay beliefs about intelligence in educational contexts,

and then extended to the study of interpersonal perception. It is important to remember that, by bridging the study of mindsets about malleability and intergroup relations, we have learned that lay beliefs matter for stereotyping, prejudice, and intergroup conflict – this was by no means an obvious or necessary extension of lay beliefs about malleability given that they are seemingly unrelated to intergroup dynamics on the surface. Yet, in the course of this research, multiple novel domains of lay theories about malleability have been identified, including beliefs about prejudice, groups, and even the world. The unique challenges of intergroup dynamics have also helped the literature on implicit theories to more fully grasp the diversity of outcomes that mindsets can influence, ranging from individual outcomes such as belonging and performance, to group outcomes such as the confrontation of prejudice and intergroup reconciliatory actions.

Furthermore, decades of research points to a simple and consistent pattern of growth mindsets yielding greater benefits than fixed mindsets, in the domains of achievement and person perception (Dweck & Leggett, 1988; Yeager et al., 2013, 2011). Yet, Rattan and Dweck (2016) have started to document conditions under which the benefits of a growth mindset may be undercut, such as in this case when targets of prejudice remain silent in the face of an expression of bias. Mendoza-Denton et al. (2008) even documented benefits of a fixed mindset, under certain conditions. Understanding the boundary conditions of the documented benefits and costs of growth and fixed mindsets, respectively, through the lens of intergroup dynamics thus illuminates and extends our understanding of the implicit theories about malleability. Further investigation may even suggest cases or conditions under which growth mindsets may prove to be maladaptive. Again, future research will have much to offer in exploring these possibilities, and will further showcase the ways in which the study of intergroup relations adds to our understandings of lay theories.

In the course of reviewing the collected evidence that mindsets influence intergroup relations, this chapter also highlights the need for more research in this domain. Much is needed, and here we highlight only a few specific areas that are most ripe for further work. There is a particular need for more investigations of when mindsets either influence or are irrelevant for social categorization, the expression of prejudice and discrimination. On the other side, targets of prejudice respond in a myriad of ways to being stigmatized, and more work could be done to investigate the role of mindsets in determining negatively-stereotyped individuals' stress and coping responses to bias, both in the short and long term. More practically, turning to real-world protracted conflicts, more research could be done to investigate how to spark and then maintain over time the sense of outgroups as malleable, given the positive trajectories that this belief may set people on.

In closing, we again highlight that the world is rife with unwanted intergroup bias. However, we suggest that the landscape of intergroup bias, and targets' responses to it, is not one size fits all. Rather, there is systematic variation in intergroup dynamics shaped by people's mindsets about malleability. Approaching the study of intergroup relations with an understanding of mindsets about malleability will offer greater insights and deeper understandings to the field of

psychology. Because people's mindsets can be changed, this approach also offers untold potential for improving intergroup harmony and equity in society, if this is the ultimate goal. We look forward to the future research that follows from an approach to intergroup relations that considers people's lay theories about malleability.

References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*(2), 113–125. doi:10.1006/jesp.2001.1491
- Bal, M., & Van den Bos, K. (2017). Effects of lay beliefs on the justice motive. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Bar-Tal, D. (2001). Why does fear override hope in societies engulfed by intractable conflict, as it does in the Israeli society? *Political Psychology, 22*(3), 601–627. doi:10.1111/0162-895X.00255
- Bar-Tal, D., & Halperin, E. (2011). Socio-psychological barriers to conflict resolution. In D. Bar-Tal (Ed.), *Intergroup conflicts and their resolution: Social psychological perspective* (pp. 217–240). New York, NY: Psychology Press.
- Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology, 42*(2), 228–235. doi:10.1016/j.jesp.2005.03.003
- Bergen, R. (1991). *Beliefs about intelligence and achievement-related behaviors*. Urbana-Champaign: University of Illinois.
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246–263. doi:10.1111/j.1467-8624.2007.00995.x
- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology, 55*(5), 726–737. doi:10.1037/0022-3514.55.5.726
- Bogardus, E. (1947). Measurement of personal-group relations. *Sociometry, 10*(4), 306–311. Retrieved from <http://www.jstor.org/stable/2785570>
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues, 55*(3), 429–444. doi:10.1111/0022-4537.00126
- Butz, D. A., & Plant, E. A. (2009). Prejudice control and interracial relations: The role of motivation to respond without prejudice. *Journal of Personality, 77*(5), 1311–1342. doi:10.1111/j.1467-6494.2009.00583.x
- Cardwell, M. (1996). *Dictionary of psychology*. Chicago, IL: Fitzroy Dearborn.
- Carr, P. B., Dweck, C. S., & Pauker, K. (2012). “Prejudiced” behavior without prejudice? Beliefs about the malleability of prejudice affect interracial interactions. *Journal of Personality and Social Psychology, 103*(3), 452–471. doi:10.1037/a0028849
- Cehajic, S., Brown, R., & González, R. (2009). What do I care? Perceived ingroup responsibility and dehumanization as predictors of empathy felt for the victim group. *Group Processes & Intergroup Relations, 12*(6), 715–729. doi:10.1177/1368430209347727
- Chiu, C.-Y., Hong, Y.-Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology, 73*(1), 19–30. doi:10.1037/0022-3514.73.1.19

- Cimpian, A., Mu, Y., & Erickson, L. C. (2012). Who is good at this game? Linking an activity to a social category undermines children's achievement. *Psychological Science*, *23*(5), 533–541. doi:[10.1177/0956797611429803](https://doi.org/10.1177/0956797611429803)
- Cohen-Chen, S., Crisp, R. J., & Halperin, E. (2015). Perceptions of a changing world induce hope and promote peace in intractable conflicts. *Personality and Social Psychology Bulletin*, *41*(4), 498–512. doi:[10.1177/0146167215573210](https://doi.org/10.1177/0146167215573210)
- Cohen-Chen, S., Halperin, E., Crisp, R. J., & Gross, J. J. (2014). Hope in the Middle East: Malleability beliefs, hope, and the willingness to compromise for peace. *Social Psychological and Personality Science*, *5*(1), 67–75. doi:[10.1177/1948550613484499](https://doi.org/10.1177/1948550613484499)
- Coleman, P. T. (2003). Characteristics of protracted, intractable conflict: Toward the development of a metaframework-I. *Peace and Conflict: Journal of Peace Psychology*, *9*(1), 1–37. doi:[10.1207/S15327949PAC0901_01](https://doi.org/10.1207/S15327949PAC0901_01)
- Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*(3), 359–378. doi:[10.1037/0022-3514.82.3.359](https://doi.org/10.1037/0022-3514.82.3.359)
- Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. *Advances in Experimental Social Psychology*, *39*, 163–254. doi:[10.1016/S0065-2601\(06\)39004-1](https://doi.org/10.1016/S0065-2601(06)39004-1)
- Crisp, R. J., Husnu, S., Meleady, R., Stathi, S., & Turner, R. N. (2010). From imagery to intention: A dual route model of imagined contact effects. *European Review of Social Psychology*, *21*, 188–236. doi:[10.1080/10463283.2010.543312](https://doi.org/10.1080/10463283.2010.543312)
- Cuddy, A. J. C., Rock, M. S., & Norton, M. I. (2007). Aid in the aftermath of Hurricane Katrina: Inferences of secondary emotions and intergroup helping. *Group Processes & Intergroup Relations*, *10*(1), 107–118. doi:[10.1177/1368430207071344](https://doi.org/10.1177/1368430207071344)
- Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious dangers of benevolent sexism: Consequences for women's performance. *Journal of Personality and Social Psychology*, *93*(5), 764–779. doi:[10.1037/0022-3514.93.5.764](https://doi.org/10.1037/0022-3514.93.5.764)
- Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science*, *314*(5798), 435. doi:[10.1126/science.1131100](https://doi.org/10.1126/science.1131100)
- de Dreu, C. K. W. (2010). Social value orientation moderates ingroup love but not outgroup hate in competitive intergroup conflict. *Group Processes & Intergroup Relations*, *13*(6), 701–713. doi:[10.1177/1368430210377332](https://doi.org/10.1177/1368430210377332)
- Deitch, E. A., Barsky, A., Butz, R. M., Chan, S., Brief, A. P., & Bradley, J. C. (2003). Subtle yet significant: The existence and impact of everyday racial discrimination in the workplace. *Human Relations*, *56*(11), 1299–1324. doi:[10.1177/00187267035611002](https://doi.org/10.1177/00187267035611002)
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18. doi:[10.1037//0022-3514.56.1.5](https://doi.org/10.1037//0022-3514.56.1.5)
- Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin*, *21*(11), 1139–1150. doi:[10.1177/01461672952111002](https://doi.org/10.1177/01461672952111002)
- Dixon, K. A., Støren, D., & Van Horn, C. E. (2002). *A workplace divided: How Americans view discrimination and race on the job*. Work trends: Americans' attitudes about work, employers and government.
- Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V. M. (2013). Prejudice, stereotyping and discrimination: Theoretical and empirical overview. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 3–28). London/Thousand Oaks: SAGE Publications Ltd.
- 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. doi:[10.1006/jesp.1997.1331](https://doi.org/10.1006/jesp.1997.1331)
- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*(3), 316–326. doi:[10.1177/0146167297233009](https://doi.org/10.1177/0146167297233009)
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. New York: Psychology Press.

- Dweck, C. S., Chiu, C.-Y., & Hong, Y.-Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, 6(4), 267–285. doi:[10.1207/s15327965pli0604_1](https://doi.org/10.1207/s15327965pli0604_1)
- Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256–273. doi:[10.1037/0033-295X.95.2.256](https://doi.org/10.1037/0033-295X.95.2.256)
- Eberhardt, J. L., Dasgupta, N., & Banaszynski, T. L. (2003). Believing is seeing: The effects of racial labels and implicit beliefs on face perception. *Personality and Social Psychology Bulletin*, 29(3), 360–370. doi:[10.1177/0146167202250215](https://doi.org/10.1177/0146167202250215)
- Ely, R. J., Meyerson, D. E., & Davidson, M. N. (2006). Rethinking political correctness. *Harvard Business Review*, 84(9), 78.
- Emerson, K. T. U., & Murphy, M. C. (2015). A company I can trust? Organizational lay theories moderate stereotype threat for women. *Personality and Social Psychology Bulletin*, 41(2), 295–307. doi:[10.1177/0146167214564969](https://doi.org/10.1177/0146167214564969)
- Erdley, C. A., & Dweck, C. S. (1993). Children's implicit personality theories as predictors of their social judgments. *Child Development*, 64(3), 863–878. doi:[10.1111/j.1467-8624.1993.tb02948.x](https://doi.org/10.1111/j.1467-8624.1993.tb02948.x)
- Feagin, J. R., & Sikes, M. P. (1994). *Living with racism: The black middle-class experience*. Boston: Beacon Press.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902. doi:[10.1037/0022-3514.82.6.878](https://doi.org/10.1037/0022-3514.82.6.878)
- 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. doi:[10.1016/S0065-2601\(08\)60317-2](https://doi.org/10.1016/S0065-2601(08)60317-2)
- Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, 20(10), 1183–1188. doi:[10.1111/j.1467-9280.2009.02422.x](https://doi.org/10.1111/j.1467-9280.2009.02422.x)
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118(2), 247–279. doi:[10.1037/a0022327](https://doi.org/10.1037/a0022327)
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York: Simon & Schuster.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24(6), 645–662. doi:[10.1016/j.appdev.2003.09.002](https://doi.org/10.1016/j.appdev.2003.09.002)
- Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, 102(4), 700–717. doi:[10.1037/a0026659](https://doi.org/10.1037/a0026659)
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27. doi:[10.1037/0033-295X.102.1.4](https://doi.org/10.1037/0033-295X.102.1.4)
- Halevy, N., Bornstein, G., & Sagiv, L. (2008). “In-group love” and “out-group hate” as motives for individual participation in intergroup conflict: A new game paradigm. *Psychological Science*, 19(4), 405–411. doi:[10.1111/j.1467-9280.2008.02100.x](https://doi.org/10.1111/j.1467-9280.2008.02100.x)
- Halevy, N., Weisel, O., & Bornstein, G. (2011). “In-group love” and “out-group hate” in repeated interaction between groups. *Journal of Behavioral Decision Making*, 25(2), 188–195. doi:[10.1002/bdm.726](https://doi.org/10.1002/bdm.726)
- Halperin, E., Crisp, R. J., Husnu, S., Trzesniewski, K. H., Dweck, C. S., & Gross, J. J. (2012). Promoting intergroup contact by changing beliefs: Group malleability, intergroup anxiety, and contact motivation. *Emotion*, 12(6), 1192–1195. doi:[10.1037/a0028620](https://doi.org/10.1037/a0028620)
- Halperin, E., Russell, A. G., Trzesniewski, K. H., Gross, J. J., & Dweck, C. S. (2011). Promoting the Middle East peace process by changing beliefs about group malleability. *Science*, 333(6050), 1767–1769. doi:[10.1126/science.1202925](https://doi.org/10.1126/science.1202925)

- Harrell, J. P., Hall, S., & Taliaferro, J. (2003). Physiological responses to racism and discrimination: An assessment of the evidence. *American Journal of Public Health, 93*(2), 243–248. doi:[10.2105/AJPH.93.2.243](https://doi.org/10.2105/AJPH.93.2.243)
- Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science, 17*(10), 847–853. doi:[10.1111/j.1467-9280.2006.01793.x](https://doi.org/10.1111/j.1467-9280.2006.01793.x)
- Haslam, N. (2017). The origins of lay theories: The case of essentialist beliefs. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Haslam, N., Bastian, B., Bain, P., & Kashima, Y. (2006). Psychological essentialism, implicit theories, and intergroup relations. *Group Processes & Intergroup Relations, 9*(1), 63–76. doi:[10.1177/1368430206059861](https://doi.org/10.1177/1368430206059861)
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113–127. doi:[10.1348/014466600164363](https://doi.org/10.1348/014466600164363)
- Hong, Y.-Y., Chiu, C.-Y., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*(3), 588–599. doi:[10.1037/0022-3514.77.3.588](https://doi.org/10.1037/0022-3514.77.3.588)
- Hornstein, H. A. (1978). Promotive tension and prosocial behavior: A Lewinian analysis. In *Altruism, sympathy, and helping: Psychological and sociological principles* (pp. 177–207).
- Hoyt, C. L., & Burnette, J. L. (2013). Gender bias in leader evaluations: Merging implicit theories and role congruity perspectives. *Personality and Social Psychology Bulletin, 39*(10), 1306–1319. doi:[10.1177/0146167213493643](https://doi.org/10.1177/0146167213493643)
- Ito, T. A., Thompson, E., & Cacioppo, J. T. (2004). Tracking the timecourse of social perception: The effects of racial cues on event-related brain potentials. *Personality and Social Psychology Bulletin, 30*(10), 1267–1280. Retrieved from <http://psp.sagepub.com/cgi/doi/10.1177/0146167204264335>
- Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*(4), 616–626. doi:[10.1037/0022-3514.85.4.616](https://doi.org/10.1037/0022-3514.85.4.616)
- Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology, 93*(4), 544–564. doi:[10.1037/0022-3514.93.4.544](https://doi.org/10.1037/0022-3514.93.4.544)
- Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion—Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science, 21*(11), 1686–1693. doi:[10.1177/0956797610384745](https://doi.org/10.1177/0956797610384745)
- Karafantis, D. M., & Levy, S. R. (2004). The role of children's lay theories about the malleability of human attributes in beliefs about and volunteering for disadvantaged groups. *Child Development, 75*(1), 236–250. doi:[10.1111/j.1467-8624.2004.00666.x](https://doi.org/10.1111/j.1467-8624.2004.00666.x)
- Kawakami, K., Dion, K. L., & Dovidio, J. F. (1998). Racial prejudice and stereotype activation. *Personality and Social Psychology Bulletin, 24*(4), 407–416. doi:[10.1177/0146167298244007](https://doi.org/10.1177/0146167298244007)
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*(4), 686–702. <https://doi.org/10.1037/0022-3514.88.4.686>
- Krysan, M. (2000). Prejudice, politics, and public opinion: Understanding the sources of racial policy attitudes. *Annual Review of Sociology, 26*, 135–168. Retrieved from <http://www.jstor.org/stable/223440>
- Lazare, A. (2004). *On apology*. New York, NY: Oxford University Press.
- Lerner, M. J. (1980). The belief in a just world. In *The belief in a just world* (pp. 9–30). Boston, MA: Springer. doi:[10.1007/978-1-4899-0448-5_2](https://doi.org/10.1007/978-1-4899-0448-5_2)
- Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science, 347*(6219), 262–265. doi:[10.1126/science.1261375](https://doi.org/10.1126/science.1261375)

- Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, 31(4), 443–453. doi:10.1177/0146167204271651
- Levontin, L., Halperin, E., & Dweck, C. S. (2013). Implicit theories block negative attributions about a longstanding adversary: The case of Israelis and Arabs. *Journal of Experimental Social Psychology*, 49(4), 670–675. doi:10.1016/j.jesp.2013.02.002
- Levy, S. R., Chiu, C.-Y., & Hong, Y.-Y. (2006a). Lay theories and intergroup relations. *Group Processes & Intergroup Relations*, 9(1), 5–24. doi:10.1177/1368430206059855
- Levy, S. R., & Dweck, C. S. (1999). The impact of children's static versus dynamic conceptions of people on stereotype formation. *Child Development*, 70(5), 1163–1180. doi:10.1111/1467-8624.00085
- Levy, S. R., Plaks, J. E., Hong, Y.-Y., Chiu, C.-Y., & Dweck, C. S. (2001). Static versus dynamic theories and the perception of groups: Different routes to different destinations. *Personality and Social Psychology Review*, 5(2), 156–168. doi:10.1207/S15327957PSPR0502_6
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74(6), 1421–1436. doi:10.1037/0022-3514.74.6.1421
- Levy, S. R., West, T. L., Ramirez, L., & Karafantis, D. M. (2006b). The protestant work ethic: A lay theory with dual intergroup implications. *Group Processes & Intergroup Relations*, 9(1), 95–115. doi:10.1177/1368430206059874
- Livingston, R. W., & Drwecki, B. B. (2007). Why are some individuals not racially biased? Susceptibility to affective conditioning predicts nonprejudice toward Blacks. *Psychological Science*, 18(9), 816–823. doi:10.1111/j.1467-9280.2007.01985.x
- Maass, A., Salvi, D., Arcuri, L., & Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, 57(6), 981–993. doi:10.1037/0022-3514.57.6.981
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, 51(1), 93–120. doi:10.1146/annurev.psych.51.1.93
- Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology*, 69(3), 397–407. doi:10.1037/0022-3514.69.3.397
- Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology*, 23(1), 77–87. doi:10.1002/ejsp.2420230107
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism*. (pp. 91–125). Orlando, FL: Academic Press.
- Mendoza-Denton, R., Kahn, K., & Chan, W. (2008). Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology*, 44(4), 1187–1193. doi:10.1016/j.jesp.2008.03.005
- Migacheva, K., & Tropp, L. R. (2012). Learning orientation as a predictor of positive intergroup contact. *Group Processes & Intergroup Relations*, 16(4), 426–444. doi:10.1177/1368430212455854
- Moeschberger, S. L., Dixon, D. N., Niens, U., & Cairns, E. (2005). Forgiveness in Northern Ireland: A model for peace in the midst of the “troubles”. *Peace and Conflict: Journal of Peace Psychology*, 11(2), 199–214. doi:10.1207/s15327949pac1102_5
- Neel, R., & Shapiro, J. R. (2012). Is racial bias malleable? Whites' lay theories of racial bias predict divergent strategies for interracial interactions. *Journal of Personality and Social Psychology*, 103(1), 101–120. doi:10.1037/a0028237
- Pauker, K., Ambady, N., & Apfelbaum, E. P. (2010). Race salience and essentialist thinking in racial stereotype development. *Child Development*, 81(6), 1799–1813. doi:10.1111/j.1467-8624.2010.01511.x
- Pettigrew, T. F. (1959). Regional differences in anti-Negro prejudice. *Journal of Abnormal Psychology*, 59(1), 28.

- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*(5), 751–783. doi:[10.1037/0022-3514.90.5.751](https://doi.org/10.1037/0022-3514.90.5.751)
- Plaks, J. E., Levy, S. R., & Dweck, C. S. (2009). Lay theories of personality: Cornerstones of meaning in social cognition. *Social and Personality Psychology Compass, 3*(6), 1069–1081. doi:[10.1111/j.1751-9004.2009.00222.x](https://doi.org/10.1111/j.1751-9004.2009.00222.x)
- Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology, 80*(6), 876–893. doi:[10.1037/0022-3514.80.6.876](https://doi.org/10.1037/0022-3514.80.6.876)
- Plaut, V. C., Thomas, K. M., & Goren, M. J. (2009). Is multiculturalism or color blindness better for minorities? *Psychological Science, 20*(4), 444–446. doi:[10.1111/j.1467-9280.2009.02318.x](https://doi.org/10.1111/j.1467-9280.2009.02318.x)
- Rattan, A., & Ambady, N. (2013). Diversity ideologies and intergroup relations: An examination of colorblindness and multiculturalism. *European Journal of Social Psychology, 43*(1), 12–21. doi:[10.1002/ejsp.1892](https://doi.org/10.1002/ejsp.1892)
- Rattan, A., & Dweck, C. S. (2010). Who confronts prejudice? The role of implicit theories in the motivation to confront prejudice. *Psychological Science, 21*(7), 952–959. doi:[10.1177/0956797610374740](https://doi.org/10.1177/0956797610374740)
- Rattan, A., & Dweck, C. S. (2016). *What happens after prejudice is experienced in the workplace?*. How mindsets affect minorities' and women's outlook on future social relations: Manuscript submitted for publication.
- Rattan, A., Savani, K., Naidu, N. V. R., & Dweck, C. S. (2012). Can everyone become highly intelligent? Cultural differences in and societal consequences of beliefs about the universal potential for intelligence. *Journal of Personality and Social Psychology, 103*(5), 787–803. doi:[10.1037/a0029263](https://doi.org/10.1037/a0029263)
- Richeson, J. A., & Nussbaum, R. J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology, 40*(3), 417–423. doi:[10.1016/j.jesp.2003.09.002](https://doi.org/10.1016/j.jesp.2003.09.002)
- Richeson, J. A., & Shelton, J. N. (2007). Negotiating interracial interactions: Costs, consequences, and possibilities. *Current Directions in Psychological Science, 16*(6), 316–320. doi:[10.1111/j.1467-8721.2007.00528.x](https://doi.org/10.1111/j.1467-8721.2007.00528.x)
- Richeson, J. A., & Trawalter, S. (2005). Why do interracial interactions impair executive function? A resource depletion account. *Journal of Personality and Social Psychology, 88*(6), 934–947. doi:[10.1037/0022-3514.88.6.934](https://doi.org/10.1037/0022-3514.88.6.934)
- Rosenthal, L., & Levy, S. R. (2012). The relation between polyculturalism and intergroup attitudes among racially and ethnically diverse adults. *Cultural Diversity and Ethnic Minority Psychology, 18*(1), 1–16. doi:[10.1037/a0026490](https://doi.org/10.1037/a0026490)
- Rosenthal, L., Levy, S. R., & Moyer, A. (2011). Protestant work ethic's relation to intergroup and policy attitudes: A meta-analytic review. *European Journal of Social Psychology, 41*(7), 874–885. doi:[10.1002/ejsp.832](https://doi.org/10.1002/ejsp.832)
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). San Diego, CA: Academic Press.
- Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*(11), 1315–1328. doi:[10.1177/0146167200263001](https://doi.org/10.1177/0146167200263001)
- Rydell, R. J., Hugenberg, K., Ray, D., & Mackie, D. M. (2007). Implicit theories about groups and stereotyping: The role of group entitativity. *Personality and Social Psychology Bulletin, 33*, 549–558.
- Saguy, T., & Halperin, E. (2014). Exposure to outgroup members criticizing their own group facilitates intergroup openness. *Personality and Social Psychology Bulletin, 40*(6), 791–802. doi:[10.1177/0146167214525475](https://doi.org/10.1177/0146167214525475)
- Sanchez, D. T., Young, D. M., & Pauker, K. (2015). Exposure to racial ambiguity influences lay theories of race. *Social Psychological and Personality Science, 6*(4), 382–390. <https://doi.org/10.1177/1948550614562844>

- Sawyer, P. J., Major, B., Casad, B. J., Townsend, S. S. M., & Mendes, W. B. (2012). Discrimination and the stress response: Psychological and physiological consequences of anticipating prejudice in interethnic interactions. *American Journal of Public Health, 102*(5), 1020–1026. doi:[10.2105/AJPH.2011.300620](https://doi.org/10.2105/AJPH.2011.300620)
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 440–452. doi:[10.1037/0022-3514.85.3.440](https://doi.org/10.1037/0022-3514.85.3.440)
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 336–356. doi:[10.1037/0033-295X.115.2.336](https://doi.org/10.1037/0033-295X.115.2.336)
- Shelton, J. N., Richeson, J. A., Salvatore, J., & Hill, D. M. (2006). Silence is not golden: The intrapersonal consequences of not confronting prejudice. In S. Levin & C. Van Laar (Eds.), *Stigma and group inequality: Social psychological perspectives* (pp. 65–81). New York: Psychology Press. doi:[10.4324/9781410617057](https://doi.org/10.4324/9781410617057)
- Shelton, J. N., Richeson, J. A., & Vorauer, J. D. (2006b). Threatened identities and interethnic interactions. *European Review of Social Psychology, 17*(1), 321–358. doi:[10.1080/10463280601095240](https://doi.org/10.1080/10463280601095240)
- Simon, S., & Hoyt, C. L. (2008). Exploring the gender gap in support for a woman for president. *Analyses of Social Issues and Public Policy, 8*(1), 157–181. doi:[10.1111/j.1530-2415.2008.00167.x](https://doi.org/10.1111/j.1530-2415.2008.00167.x)
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811. doi:[10.1037//0022-3514.69.5.797](https://doi.org/10.1037//0022-3514.69.5.797)
- Sue, D. W. (2010). *Microaggressions in everyday life: Race, gender, and sexual orientation*. New York: Wiley.
- Swim, J. K., & Hyers, L. L. (1999). Excuse me—What did you just say?!: Women’s public and private responses to sexist remarks. *Journal of Experimental Social Psychology, 35*(1), 68–88. doi:[10.1006/jesp.1998.1370](https://doi.org/10.1006/jesp.1998.1370)
- Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. *Journal of Social Issues, 57*(1), 31–53. doi:[10.1111/0022-4537.00200](https://doi.org/10.1111/0022-4537.00200)
- Swim, J. K., Hyers, L. L., Cohen, L. L., Fitzgerald, D. C., & Bylsma, W. H. (2003). African American college students’ experiences with everyday racism: Characteristics of and responses to these incidents. *Journal Of Black Psychology, 29*(1), 38–67. doi:[10.1177/0095798402239228](https://doi.org/10.1177/0095798402239228)
- Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 83–114). L. Erlbaum Associates.
- Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin, 37*(8), 1055–1067. doi:[10.1177/0146167211406506](https://doi.org/10.1177/0146167211406506)
- Tropp, L. R., & Pettigrew, T. F. (2005a). Differential relationships between intergroup contact and affective and cognitive dimensions of prejudice. *Personality and Social Psychology Bulletin, 31*(8), 1145–1158. doi:[10.1177/0146167205274854](https://doi.org/10.1177/0146167205274854)
- Tropp, L. R., & Pettigrew, T. F. (2005b). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science, 16*(12), 951–957. doi:[10.1111/j.1467-9280.2005.01643.x](https://doi.org/10.1111/j.1467-9280.2005.01643.x)
- Vallacher, R. R., Coleman, P. T., Nowak, A., & Bui-Wrzosinska, L. (2010). Dynamical foundations of intractable conflict: Introduction to the special issue. *Peace and Conflict: Journal of Peace Psychology, 16*(2), 113–125. doi:[10.1080/10781911003691294](https://doi.org/10.1080/10781911003691294)
- Vescio, T. K., Gervais, S. J., Snyder, M., & Hoover, A. (2005). Power and the creation of patronizing environments: The stereotype-based behaviors of the powerful and their effects on female performance in masculine domains. *Journal of Personality and Social Psychology, 88*(4), 658–672. doi:[10.1037/0022-3514.88.4.658](https://doi.org/10.1037/0022-3514.88.4.658)

- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*(5), 961–977. doi:[10.1037/0022-3514.45.5.961](https://doi.org/10.1037/0022-3514.45.5.961)
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*(4), 548–573. doi:[10.1037/0033-295X.92.4.548](https://doi.org/10.1037/0033-295X.92.4.548)
- Weiner, B., Heckhausen, H., & Meyer, W.-U. (1972). Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. *Journal of Personality and Social Psychology*, *21*(2), 239–248. doi:[10.1037/h0032307](https://doi.org/10.1037/h0032307)
- Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, *94*(6), 1033–1047. <https://doi.org/10.1037/0022-3514.94.6.1033>
- Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/ethnic discrimination and health: Findings from community studies. *American Journal of Public Health*, *93*(2), 200–208.
- Williams, D. R., & Williams-Morris, R. (2000). Racism and mental health: The African American experience. *Ethnicity & Health*, *5*(3–4), 243–268. doi:[10.1080/713667453](https://doi.org/10.1080/713667453)
- Wohl, M. J. A., Cohen-Chen, S., Halperin, E., Caouette, J., Hayes, N., & Hornsey, M. J. (2015). Belief in the malleability of groups strengthens the tenuous link between a collective apology and intergroup forgiveness. *Personality and Social Psychology Bulletin*, *41*(5), 714–725. doi:[10.1177/0146167215576721](https://doi.org/10.1177/0146167215576721)
- Wolfe, C. T., & Spencer, S. J. (1996). Stereotypes and prejudice. *American Behavioral Scientist*, *40*(2), 176–185.
- Woodzicka, J. A., & LaFrance, M. (2001). Real versus imagined gender harassment. *Journal of Social Issues*, *57*(1), 15–30. doi:[10.1111/0022-4537.00199](https://doi.org/10.1111/0022-4537.00199)
- Yeager, D. S., Miu, A. S., Powers, J., & Dweck, C. S. (2013). Implicit theories of personality and attributions of hostile intent: A meta-analysis, an experiment, and a longitudinal intervention. *Child Development*, *84*(5), 1651–1667. doi:[10.1111/cdev.12062](https://doi.org/10.1111/cdev.12062)
- Yeager, D. S., Trzesniewski, K. H., Tirri, K., Nokelainen, P., & Dweck, C. S. (2011). Adolescents' implicit theories predict desire for vengeance after peer conflicts: Correlational and experimental evidence. *Developmental Psychology*, *47*(4), 1090–1107. doi:[10.1037/a0023769](https://doi.org/10.1037/a0023769)
- Yzerbyt, V., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialistic view of group perception. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 20–50). Malden, MA: Blackwell Publishing.

Effects of Lay Beliefs on the Justice Motive

Michèle Bal and Kees van den Bos

Injustice anywhere is a threat to justice everywhere
Martin Luther King, Jr., (1963)

That justice is important to people seems to be an incontestable notion. Rules of justice lie at the heart of modern societies with a society's constitutional law usually defining important basic human rights and responsibilities and elaborate systems of laws and regulations guiding people's lives in the social world. In addition, individuals are greatly concerned with justice (Folger, 1984) and feel threatened by injustice, as the 1963 quote by Martin Luther King Jr. illustrates. Not only do people want fair outcomes for everyone (see, e.g., Adams, 1965; Walster, Walster, & Berscheid, 1978), but people also greatly value being treated fairly and treating each other with respect (Lind & Tyler, 1988; Tyler & Lind, 1992). So, while not many people will argue with the fact that justice is an important social value, the question of how people define justice is much more difficult to address. This issue has intrigued philosophers, legal scholars, and social scientists alike.

In general, philosophical questions related to defining justice mostly revolve around issues of what constitutes a just society and how a just world can be achieved (e.g., Rawls, 1971) or around how people can live a moral and virtuous life (e.g., Beauchamp, 2001). For instance, Rawls (1971) used the notion of a veil of ignorance, a thought experiment in which no one knows their position in society, to come to pure moral reasoning regarding the rules of justice in societal and political decision-making. With some notable exceptions, legal scholars are mostly concerned with "black-letter law," which refers to the law as it is written in legal codes and enacted by legislators (Finkel, 2000). As such, many legal scholars study how

M. Bal (✉)

Department of Interdisciplinary Social Science, Utrecht University,
Heidelberglaan 1, 3584 CS Utrecht, The Netherlands
e-mail: m.bal@uu.nl

K. van den Bos

Department of Psychology and School of Law, Utrecht University,
Heidelberglaan 1, 3584 CS Utrecht, The Netherlands

laws and legislation should work and how these may be improved legislation. “Black-letter lawyers” do not focus not on lay people’s perceptions of justice directly. In other words, both philosophical and legal perspectives on justice focus mostly on normative aspects of justice or on the so-called “ought”-questions.

In the current chapter we do not focus on philosophical, legal, or other normative conceptions of justice, but instead we will elaborate on theories of justice that lay people have. For that purpose, we take a social scientific approach to studying justice (Cohen, 1986). Building on this perspective we will elaborate on commonsense notions of justice (Finkel, 2000; Tyler, 2006) and the effects that these commonsense notions can have on people’s reactions when they have been confronted with injustice.

Commonsense justice reflects what ordinary people think is just and fair. These perceptions of justice will to a large degree overlap with philosophical and legal notions of justice. However, this may not always be the case, for instance when people protest against certain laws and regulations. A classic example in this respect is the experience of Rosa Parks who in 1955 refused to give up her bus seat to a white person and was arrested for it. This became an important event in the civil rights movement in the U.S. in the 1950s and 1960s. We believe that while lay people’s conception of justice will not always follow normative ideas of justice, perceptions of justice and injustice are of crucial importance when we want to understand how people will respond and behave in our world. After all, if men or women define situations as real, they are real in their consequences (Thomas & Thomas, 1928). We will now first review some of the classical distinctions made in social justice research.

Within social scientific theorizing on justice, a distinction is often made between distributive justice and procedural justice. The former focuses on the fairness of distributions of goods and resources. Put differently, distributive justice concerns the fairness of outcome distributions (e.g., Adams, 1965; Blau, 1964). In early theorizing researchers focused mostly on issues of distributive justice and proposed equity as an important determinant of outcome fairness judgments (Adams, 1965; Walster et al., 1978). Equity theory proposes that people prefer equal outcomes for equal inputs. More precisely, people are assumed to judge an outcome as just or fair when their own outcome-to-input ratio equals some comparative or referent outcome-to-input ratio. Several studies have shown that people dislike inequitable underpayment as well as inequitable overpayment (e.g., Adams, 1965; Peters, 2006), lending support to important predictions from equity theory.

In later studies, procedural justice was introduced as an important other determinant of people’s justice judgments. Procedural justice entails the fairness of how people arrive at certain outcomes and not on the outcomes itself (e.g., Lind & Tyler, 1988; Van den Bos, 2005, 2015). Hence, procedural justice is focused on the fairness of decision processes or (more generally) the fairness of how people are treated (Van den Bos, 2015). Procedural justice even has been proposed to be more important for understanding people’s reactions than distributive justice (Lind & Tyler, 1988, p. 1); a proposition that has gained support in several studies (see, e.g., Tyler, 1987, 1989).

We want to emphasize that the discussion above on the different types of justice is far from complete. We further note that, in addition to distinguishing distributive and procedural justice, a further distinction has been made between retributive justice (e.g., Wenzel & Mummendey, 1996; Wenzel & Okimoto, 2016) and restorative justice (e.g., Cohen, 2016). Retributive justice focuses mainly on punishment for perpetrators while restorative justice tends to concentrate on the victim's perspective. Specifically, restorative justice is concerned with re-establishing the relations between the victim, offender, and society (Cohen, 2016).

Notwithstanding the importance and relevance of the different types of justice discussed thus far, in the current chapter we want to focus on why people care about justice and how this shapes reactions following injustice. Most of the research and theorizing described above focused on *what* "types of justice" people care about and less on *why* people care about justice. Hence, the question of what motivates lay people to place importance on justice in their lives still remains. This question is addressed by Lerner in just-world theory (e.g., Lerner, 1977, 1980). In his seminal theory, he focused on lay people's conception of justice to explain reactions following confrontations with injustice. In the current chapter we will focus on this approach in explaining the effects of lay theories of justice on responses to social injustice.

Justice motive theory or just-world theory (Lerner, 1980) assumes the need for justice to be a fundamental human need and focuses on people's reactions following a confrontation with innocent suffering. According to the theory, lay people define justice as everyone getting what they deserve such that good things will happen to good people and bad things will only happen to bad people.

In this chapter we will first elaborate on the origins and functions of this just-world belief (Lerner, 1977, 1980). Subsequently, following the general tenet of just-world theory, we will focus on people's reactions toward victims of injustice. We will discuss the role of the belief in a just world in people's reactions following unjust events they observe and explain how the belief in a just world can paradoxically lead to victim blaming. Subsequently, processes that play a role in shaping these derogatory reactions will be discussed. Here we broadly distinguish two lines of research, one focusing on processes that occur before the unjust event has taken place and that influence the construal of an unjust event, and one discussing basic psychological processes that take place after a confrontation with an innocent victim and that influence the processing of an unjust event. These processes are illustrated in Fig. 1. Both types of processes can influence reactions toward innocent victims. In the final part of this chapter, we will also describe studies on alternative lay people's reactions to deal with unjust situations and alternative lay people's operationalizations of justice that may help explain the broad range of possible reactions following unjust events.

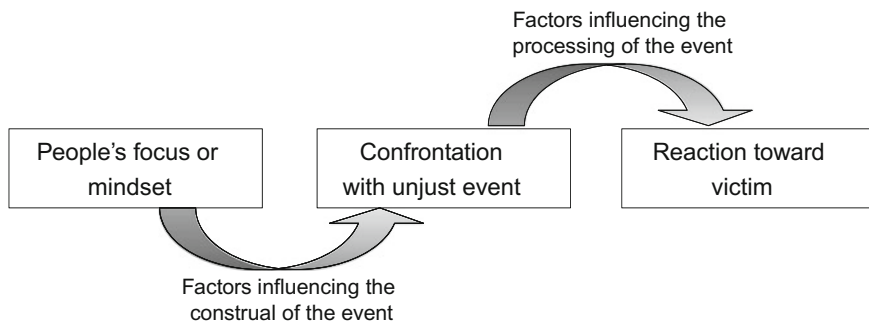


Fig. 1 Schematic overview of the factors influencing people's reactions toward unjust events

The Belief in a just World

Lerner (1977, 1980) argued that the fundamental need for a just world, that is, a world in which people get what they deserve, stems from a personal contract that is adopted in childhood when children learn to give up immediate satisfaction for more delayed—and often greater—rewards. Believing in a just world provides structure to our social world and gives people the confidence that efforts will pay off. When people do not adopt this belief, striving for delayed rewards would seem futile, as they cannot be certain that their efforts will pay off. As such, the belief in a just world enables people to focus on the future, strive for long-term goals and trust that their efforts will pay off in the end. Studies have shown that, indeed, people defend their belief in a just world more vigorously when they are focused on the future as opposed to the present (e.g., Bal & Van den Bos, 2012; Hafer, 2000a; Hafer & Rubel, 2015; Laurin, Fitzsimons, & Kay, 2011). Moreover, research shows that this effect is due to feelings of uncertainty being reduced by endorsing the belief in a just world (Bal & Van den Bos, 2012).

While the main function of the belief in a just world is making the world predictable and enabling people to focus on the future (Bal & Van den Bos, 2012; Hafer, 2000a), believing in a just world has also been related to several psychological health indices. For this purpose, several researchers have constructed scales to measure the degree to which people believe in a just world (e.g., Lipkus, Dalbert, & Siegler, 1996) and related these to various outcome measures. In general, studies found that the more people believe that the world is just, the higher their well-being, positive affect, optimism, and the more effectively they can cope with stress (for overviews, see Furnham, 2003; Hafer & Sutton, 2016). These studies show that it generally seems adaptive to believe that the world is a just place.

Moreover, these studies also showed that people can differ in the strength with which they endorse the belief in a just world. That is, there may be individual and cultural differences in the strength with which people express their belief in a just world (see, e.g., Sabbagh & Schmitt, 2016; Schmitt, Baumert, Gollwitzer, & Maes, 2010). However, we argue that it is important to distinguish between people's

tendency to endorse this belief, which can be related to several health and other indices, and people's general need for a just world and the related threat experienced by injustice, which seems to be universal (Hafer & Sutton, 2016). Hence, while people can differ in how strongly they express the belief in a just world, all people will likely still experience a sense of threat when confronted with injustice. In this chapter, we will primarily focus on this more or less universal need for a just world and subsequent reactions to deal with unjust events.

On a daily basis we are confronted with many instances of unjust events. When we watch the news, read a newspaper, or talk to friends or acquaintances, we often see or hear stories of injustice and unjust events that happen to people in our world. These stories can encompass minor instances of injustice, such as catching a bad break, but also more grave unjust situations, such as discrimination, terrorist acts, serious crimes or other types of violence in which innocent people are victimized. How do people maintain their belief in a just world in the face of such great evidence to the contrary?

These unjust situations should make it impossible or at least very difficult to uphold the belief in a just world. Yet, because this belief serves so many important functions for individuals, it is also impossible to give up on it when an unjust event is encountered. People are not ignorant about these unjust events and they do not deny that injustice exists in the world in general. However, people do maintain that the world is just for them personally as it serves such important social functions. Hence, people make a distinction between the world in general and their personal world (Lerner, 1980; Lipkus et al., 1996; Sutton & Douglas, 2005). This personal world does not only encompass them as individuals, but consists of that part of the world in which they live and function. The scope of this world, that is, who is or is not included in this personal world, may differ as a function of the situation (Lerner, 1980; Opatow, 1990). Differentiating between a personal world and the world in general enables people to uphold the belief in a personally just world while at the same time acknowledging that injustice does exist in the world in general.

When an unjust situation concerns people's personal world directly, they can react to unjust suffering in various ways to restore their belief in a just world. First, experienced or perceived injustice can evoke strong emotional reactions. Moral outrage has been coined as a specific negative emotion following acts of injustice (Montada & Schneider, 1989). Moreover, while reactions to grave unjust situations are to be expected, even minor events can instill a sense of injustice (Gaucher, Hafer, Kay, & Davidenko, 2010).

Importantly, reactions to deal with a confrontation with unjust suffering can diverge greatly. Sometimes people will stand up against injustice and go to great lengths to "right a wrong," for instance by punishing the perpetrators or by compensating the victims. Sometimes people even March the streets to protest against grave unjust situations. At other times, however, people tend to blame victims for their ill plight, stating that these victims must have done something to deserve what happened to them. Paradoxically, both types of reactions can be explained by lay people's "concern for justice" (Lerner, 1980). That is, standing up against unjust situations, helping victims, and punishing perpetrators are all ways in which people

can actively pursue a more just world. On the other hand, blaming the victims also helps to restore one's sense of justice, as the victims become deserving of their ill fates and, hence, the unjust situation is cognitively resolved.

In just-world theory, Lerner (1980) included both active strategies, such as helping or compensating the victim, as well as more passive or cognitive strategies, such as victim blaming, as ways to resolve a threat to people's just-world beliefs. Over the past decades, research on reactions following unjust events has accumulated. Most of these studies, however, focused on victim blaming. In the following section we will discuss some important studies that have been conducted in this area and that provide insight into the processes that are involved in victim blaming predominantly, but also other strategies, to uphold the belief in a just world.

Psychological Processes Underlying Victim Blaming

Most of the research on victim blaming focused on situational factors that either increased or decreased the just-world threat experienced, resulting in more victim blaming and derogation, or decreased the just-world threat, yielding less blaming and derogation of innocent victims. Whereas blaming is focused on condemning a victim's actions, derogation is focused on condemning a victim's character. These studies showed, for instance, that people blamed a victim more when the victim's suffering was enduring as opposed to ending (e.g., Hafer, 2000a, Study 2) and when the victim actually did something to contribute to the injustice occurring (i.e., a non-innocent victim; e.g., Hafer, 2000a, Study 1). Moreover, people also blamed a victim more when the perpetrator was not caught as opposed to caught, presumably because chances of justice being served increase when the perpetrator has been apprehended (e.g., Hafer, 2000b; Van Prooijen & Van den Bos, 2009; Van den Bos & Maas, 2009). Intuitively, these findings make sense as they can easily be explained by the fact that a greater just-world threat would lead to more victim blaming.

In a seminal paper, Hafer (2000b) showed that these effects were indeed due to an increased activation of justice-related constructs. That is, in two studies she confronted participants with a scenario in which a boy was severely assaulted and robbed. In these studies, Hafer used a manipulation of perpetrator apprehension. That is, half of the participants were told that the perpetrator had been caught and was sent to jail, creating a low threat, while the other half of the participants read that the perpetrator was still at large and would not likely be caught, creating a high threat. Subsequently, participants' concern with justice was measured in an implicit manner, using a modified Stroop task. In a Stroop task words in different colors are presented to participants. Participants have to identify the color of words presented to them and ignore the content of these words. In Hafer's version of the Stroop task, justice-related, harm-related, story-related, and neutral words were included. Results revealed that people who read that the perpetrator had not been apprehended experienced took longer to identify the color of justice-related words (as

opposed to other neutral, story-related or harm-related words) than the people who read that the perpetrator had been apprehended. This can be explained by the fact that it is more difficult to ignore the content of words that are activated in your mind (in this case “justice”).

These findings indicate that people’s concern with justice is heightened following a confrontation with injustice, especially when this instance constitutes a stronger threat to the belief in a just world. In a follow-up study, Hafer also revealed that victim blaming and derogation reduced this concern, as this interference attenuated for participants who were given a chance to blame and derogate the victim as opposed to those who did not get this opportunity. Hence, this indicates that victim blaming and derogation serve as viable ways to resolve the threat to people’s just world, posed by a confrontation with innocent suffering.

In addition to victim innocence, perpetrator apprehension, and enduring versus ended victim suffering, victim similarity has been put forward as another possible variable influencing threat to the belief in a just world (e.g., Bal & Van den Bos, 2010; Correia, Vala, & Aguiar, 2001, 2007). Specifically, when a victim is more similar (versus less similar) to an observer, two hypotheses can be put forward. On the one hand, one could expect that similarity increases identification and with that empathy for the victim, which would reduce victim blaming. On the other hand, one could expect that similarity would increase the fear of a similar fate bestowing on the observer and therefore victim blaming will be enhanced.

Studies focusing on victim similarity showed support for the latter hypothesis. When a victim belonged to the same social group as the observer, victim blaming was increased (Correia et al., 2001, 2007). Later studies added to these findings by showing a similar effect for perpetrator similarity. That is, belonging to the same social group as the victim or to the same social group as the perpetrator both increased victim blaming (Bal & Van den Bos, 2010). Moreover, social similarity to a victim or perpetrator (i.e., belonging to the same social group) as well as physical proximity of an unjust event both increase negative reactions toward a victim (e.g., Bal & Van den Bos, 2012, Study 2, 2015).

More recently, studying the processes that are involved in processing injustice has become more prominent in just-world research. These studies focused on how certain psychological variables influenced lay people’s reactions following a confrontation with injustice. We will discuss this process-oriented research in more detail in the following paragraphs. In doing so, we first focus on factors that influence how people construe an unjust event and subsequently move on to a discussion of the processes that take place after people have been confronted with innocent suffering, and that influence how people process an unjust event (see Fig. 1 for a schematic overview).

A Focus on the Self Versus Others

In the 1970s, several researchers studied the assignment of blame or responsibility to victims of accidents or other types of injustice (e.g., Chaiken & Darley, 1973; Lerner & Simmons, 1966; Novak & Lerner, 1968; Shaver, 1970; Walster, 1966). These seminal studies inspired many others to further investigate these issues and led to the distinction between person identification and position identification (Lerner, Miller & Holmes, 1976). In the former, people are more concerned with the ill fate of the victim (“I feel his suffering”), while in the second, they will be more focused on their own personal consequences (“That could also happen to me”).

More recently, research on the relation between self-construal and victim blaming (Van Prooijen & Van den Bos, 2009) extended these earlier findings by researching the more indirect processes influencing the construal of unjust events. Self-construal refers to whether people describe *themselves* in terms of group membership or in terms of individual qualities (Singelis, 1994). That is, people can either define themselves in terms of differences with others, stressing their uniqueness (i.e., an independent self-construal), or in terms of similarities with others, stressing their belonging to certain social groups (i.e., an interdependent self-construal). People dispositionally and situationally differ in whether they adopt a more independent or interdependent self-construal. In their research, Van Prooijen and Van den Bos (2009) applied this insight to the study of victim blaming. In this study participants read a scenario in which a girl was assaulted after a night out, after which blaming of the victim was measured. The researchers found that both manipulated and measured high levels of *interdependent* self-construal led participants to blame the victim more than when they were primed with or scored high on *independent* self-construal. According to the Van Prooijen and Van den Bos, these findings can be explained by the fact that an interdependent self-construal may facilitate assimilation with others (i.e., position identification). When people assimilate with a victim specifically, this may enhance the threat experienced and thus increase derogatory reactions toward this victim.

In line with these findings, our research showed that victim blaming is enhanced when people are self-focused as opposed to other-focused (Bal & Van den Bos, 2015), presumably because people who are self-focused will be more concerned with the threat that a situation of injustice poses, while other-focused individuals will be more concerned with the victim’s fate. Hence, a self-focus may lead to position identification while an other-focus may lead to person identification. In our studies, half of the participants were asked to think back to and describe a situation in which they were focused on themselves (e.g., studying for an exam) and the other half to think back to and describe a situation in which they were focused on others (e.g., listening to a lecturer giving a lecture). Subsequently, we confronted the participants with a scenario in which a man was severely injured after being hit by a car. Our findings on victim blaming showed that reactions were enhanced when they were self-focused as opposed to other-focused.

Additional studies, focusing on related processes, also speak to the fact that a self-focus enhances and an other-focus reduces derogatory reactions toward victims. For instance, studies have shown that mimicking a person, whether it be the victim or a person unrelated to the situation, reduced victim blaming (Stel, Van den Bos, & Bal, 2012). According to the authors, mimicking might induce a general other-oriented mindset. Put differently, mimicking might make an other-focus or person identification more likely. Moreover, studies have shown that ego depletion (i.e., being low on self-control) enhances victim blaming and, perhaps even more important for the current discussion, that self-affirmation reduces victim blaming (Loseman & Van den Bos, 2012). Loseman and Van den Bos (2012) argue that these findings may be explained by the fact that the victim poses a self-threat to the observer. Hence, these authors again relate a self-focus to the experienced threat and subsequently to more victim blaming.

All and all, a picture emerges that fits with the idea that people may construe a just-world threat differently depending on whether they are mainly focused on personal consequences as opposed to how the victim must feel. These findings are in line with earlier theorizing (Lerner, Miller, & Holmes, 1976). Identification with the position of the victim enhances experienced threat and therefore derogatory reactions to deal with this threat, while identification with the victim as a person enhances sympathy and, as such, will decrease derogatory reactions. Some papers studying distributive and procedural justice also alluded to the fact that these focuses could influence justice judgments and related reactions following personally experienced injustice (e.g., Lerner & Clayton, 2011; Skitka, Aramovich, Lytle, & Sargis, 2009; Van Prooijen, 2013). We will now move on to a discussion of the processes that take place following such confrontation with an innocent victim.

Approach Versus Avoidance Orientation

The research described above focused on factors influencing the construal of the event, *before* people are confronted with injustice. We will now turn to a discussion of studies focusing on what happens *after* people have been confronted with an unjust situation and look for factors that may influence subsequent processing of unjust information. Hence, we will discuss the processes that take place in between the confrontation with injustice and people's overt reactions toward the victims. A seminal motivational dichotomy that influences a broad range of psychological phenomena is that of approach and avoidance (Chen & Bargh, 1999). Approach and avoidance motivation have been found to play a role in most, if not all, human behavior. That is, people will be motivated to avoid negative stimuli (punishment) as much as possible and to approach positive stimuli (rewards) when they can. Similar to a self versus other focus, this orientation to approach or avoid can differ dispositionally and situationally. Approach and avoidance motivation will likely also influence how people process a confrontation with unjust suffering. Specifically, we expect that when people are approach motivated toward victims, a

concern with their ill fate will be likely, while avoidance motivation might make more derogatory and rejecting reaction more likely. We have studied how these motivations influence reactions toward victims of injustice by experimentally inducing them before a confrontation with an innocent victim and by measuring them after such a confrontation.

With regard to the former, our research has shown that people tend to blame an innocent victim more when they are avoidance as opposed to approach motivated (Bal, 2014; Bal & Van den Bos, 2016). We further showed that a confrontation with an innocent victim who poses a high just-world threat inhibits people's natural approach tendencies and leads people to become more avoidance than approach motivated toward the victim. These findings indicate that people will oftentimes react in an avoidance-motivated manner toward a confrontation with an innocent victim. Such an avoidance-motivated reaction may heighten chances of victim blaming as a way of resolving the threat to one's belief in a just world.

Experiential Versus Rationalistic Processing

Another way to study what happens after people have been confronted with a victim is by looking at the influence of rationalistic versus experiential processing of the unjust event. According to dual-process theories (e.g., Strack & Deutsch, 2004), people can process information in one of two ways. They either use rationalistic and effortful routes, in which information is processed in detail and in which costs and benefits are carefully weighed against alternative options. Alternatively they can use experiential and intuitive routes, which process information more quickly and superficially and work by using heuristics. Following the increased attention to dual-process theories within the psychological literature (e.g., Strack & Deutsch, 2004), the influence of rationalistic and experiential processing in reactions following injustice has also gained attention in social justice research (e.g., Harvey & Callan, 2014; Van den Bos & Maas, 2009; Van den Bos et al., 2008). Within the justice motive literature there is an ongoing debate about whether rationalistic or experiential processing of the situation is dominant in people's reactions to the unjust event (e.g., Van den Bos, 2007; Lerner & Clayton, 2011). A number of researchers argue that reactions toward unjust situations result from intuitive experiential processing of information (e.g., Lerner & Clayton, 2011; Lerner & Goldberg, 1999; Harvey & Callan, 2014). In contrast, other studies show that these reactions are stronger when people have adopted rationalistic assessments of the situation (e.g., Van den Bos & Maas, 2009). The type of justice information that is processed is important in this respect (Maas & Van den Bos, 2009; Van den Bos & Maas, 2009).

In their theorizing, Lerner and Goldberg (1999) argue that justice judgments and subsequent reactions are usually arrived at through intuitive processing of the unjust situation. That is, they propose that the management of people's just-world beliefs usually takes place outside of people's consciousness. Hence, negative reactions

following unjust events are due to the associative link of the victim to the negative event and positive reactions can occur spontaneously when helping is an available option that is effortless and relatively costless (i.e., experiential processing).

Harvey, Callan, and Matthews (2014) found partial evidence for this line of reasoning. In a series of studies these authors manipulated and measured information processing style, being either intuitive and experiential or rationalistic and effortful, and measured a variety of reactions following a confrontation with a victimization scenario, which constituted either a high or low just-world threat. Findings revealed that most reactions differed based on victim innocence and victim suffering (i.e., the just-world threat manipulations), regardless of information processing mode. Therefore, the authors concluded that people's reactions to victimization, including victim blaming, occur intuitively as well as through rationalistic processing. Hence, rationalistic processing of unjust information is not a necessary prerequisite for reacting toward innocent victims.

In contrast with the above line of reasoning, and in line with the uncertainty management model (Van den Bos, 2009), Van den Bos and Maas (2009) propose that it is rationalistic as opposed to experiential processing that occurs after a confrontation with threats to the belief in a just world. These authors conducted an experiment in which they asked participants to either react to information in an intuitive and experiential or rational and deliberative manner. After this, participants read a scenario in which a woman was either robbed or sexually assaulted and the degree to which the participants blamed the victim was measured. They induced a high or low just-world threat by telling participants that the perpetrator was either caught (low threat) or not (high threat). Their results revealed that only people in a rationalistic mindset blamed a victim more when the perpetrator was still at large as opposed to when he was caught. This difference was not there for people in an intuitive mindset. Interestingly, in an intuitive mindset, blaming was generally higher than in a rationalistic mindset. The authors concluded that rationalistic processing enhances victim blaming following a high as opposed to a low just-world threat, and hence, threat-related victim blaming is the result of rationalistic as opposed to experiential processing of the unjust situation.

Future research may want to reconcile the findings by Harvey et al. (2014) with those obtained by Van den Bos and Maas (2009). For example, it might be the case that differences in reactions found by Harvey and colleagues were related to heuristics, such as the need to reduce negative affect or (lack of) care for the victim, and not to the just-world threat the victims posed. Hence, the manipulations of victim innocence and victim suffering, adopted by Harvey et al. (2014), may have enabled differing reactions for several reasons other than the just-world threat the victim posed. In contrast, participants in the studies by Van den Bos and Maas (2009) may have focusing on the just-world threat that the victim posed specifically, for which rationalistic processing seems necessary.

Interestingly, with regard to procedural justice judgments, differences were found mainly in an experiential mindset as opposed to a rationalistic mindset (Maas & Van den Bos, 2009). That is, in a set of studies in which participants reacted to a fair or unfair procedure, results showed that especially in an experiential mindset

did people react more negatively to an unfair as opposed to a fair procedure. In a rationalistic mindset no differences of procedural fairness were found. It may be the case that in these instances experiential processing fits the context better, because affect and procedural justice are linked (i.e., feeling bad because of procedural injustice). Hence, while reacting to personally experienced procedural unfairness seems to be a more experiential and intuitive process, interpreting a confrontation with an innocent victim in terms of the just-world threat that the situation poses seems to be a more rationalistic process.

Taking these results together, the studies converge and diverge on certain points with regard to the processing of information on unjust suffering. That is, while researchers on each side of the debate stress the importance of either intuitive or rationalistic processing in reactions to innocent suffering and victim blaming specifically, both sides do agree on the fact that intuitive as well as rationalistic paths to reacting to unjust situations are possible. It seems to be the case that only reactions following rationalistic, effortful and deliberative processing of innocent suffering are sensitive to threat-related information, as shown by Van den Bos and Maas (2009). Spontaneous, intuitive, and experiential reactions to unjust situations, on the other hand, are not influenced by the degree of threat that the situation poses, as suggested by Harvey et al. (2014) and Maas and Van den Bos (2009).

Evidence for Positive Reactions Following Unjust Events

Most of the research inspired by the introduction of justice motive theory (Lerner, 1980) focused on factors influencing derogatory reactions of victim blaming and derogation (for an overview, see Hafer & Bègue, 2005). However, reactions following unjust events can be much more varied, as already explained in the introduction of this chapter. That is, oftentimes people do not react in derogatory terms toward victims, but unjust situations spark strong negative emotions and a willingness to take action against the unjust event. Outside the realm of justice motive theory, it has been found that people experience moral outrage following a confrontation with injustice and go to great lengths to alleviate the victim's ill plight or punish the wrongdoer, sometimes even by sacrificing their own positive outcomes (Batson, 1998).

In just-world theory, Lerner (1980) already alluded to the possibility that people react in this more constructive way toward confrontations with injustice, actively pursuing a (more) just world. Specifically, Lerner distinguished helping and compensating the victim together with punishing the perpetrator from blaming and derogating the victim together with other more "irrational" strategies to preserve the belief in a just world. In a seminal study, Lerner and Simmons (1966) found that people will help a victim when helping is an available option, and only resort to victim blaming when helping is deemed futile.

In more recent work, attention is also given to these more positive reactions and the question of how people choose to adopt a certain strategy for resolving a

just-world threat (see, e.g., Bègue, Charmoillaux, Cochet, Cury, & De Suremain, 2008; DePalma, Madey, Tillman, & Wheeler, 1999; Hafer & Gosse, 2011; Hafer & Rubel, 2015; Kogut, 2011). By and large, these studies focused on dispositional traits influencing willingness to help and did not study the underlying processes involved in deciding how to react to an unjust event.

In our own studies, we did include helping or supporting the victim as a possible reaction in several studies and investigated the role of approach and avoidance motivation as well as a self- versus and other-focus also in relation to positive reactions. Our findings showed that while a self-focus enhanced victim blaming, an other-focus decreased victim blaming and enhanced support for the victim. That is, after presenting the participants with a car crash scenario, we measured whether people were willing to invest time and effort into raising money for the victim of the car crash. Our results showed that when people were other-focused, they helped the victim more than when they were self-focused (Bal & Van den Bos, 2015). Moreover, in a different set of studies we also found that people react more supportively and less derogatory toward the victim when they were approach motivated as opposed to avoidance motivated (Bal, 2014; Bal & Van den Bos, 2016).

In our studies, it did not seem to be the case that people necessarily help when helping was possible. Instead, oftentimes people's spontaneous reactions were to cognitively resolve a just-world threat by resorting to victim blaming and derogation. When people were explicitly made to focus on the victim's well-being (by inducing an approach motivation or an other-focus), they did opt to help more as a way of resolving a just-world threat. These findings may be reconciled with (Lerner's 1980; Lerner & Simmons, 1966) propositions by taking into account the costs of helping. When helping is relatively costless and effortless, people will help a victim. When helping involves effortful and costly behavior, for instance by having to spend time or money, people will be less willing to help and may adopt a cognitive strategy of victim blaming or derogation as a relatively more likely option.

In addition to derogatory and supportive reactions toward the victims of misfortune, more differentiated reactions are possible and have received some attention in research. We want to address two related types of reactions, namely immanent and ultimate justice reasoning (Callan, Sutton, Harvey & Dawtry, 2014; Callan, Ellard & Nicol, 2006; Harvey & Callan, 2014) and compensatory rationalizations (Kay et al., 2007; Kay & Jost, 2003; Jost & Kay, 2005; Gaucher et al., 2010). Both types of reactions try to make sense of an unjust situation by placing it in a broader perspective, assuming that justice and injustice balance out. In the former, people perceive misfortunes as caused by previous bad deeds (immanent justice reasoning) or resulting in ultimate compensation (ultimate justice reasoning). In the latter, people are expected to keep up a kind of moral balance or create an illusion of equality such that negative traits or undeserving events are "compensated" with positive traits or events with opposite valence. So in addition to people getting what they deserve, a balance between good and bad outcomes may be a lay theory that people adopt in the realm of justice.

Immanent justice reasoning can be viewed as the belief that actions bring about deserved outcomes. In this type of reasoning, people make a causal link between prior moral behavior and subsequent random outcomes. Importantly, Callan and colleagues stress that the lack of a physically plausible means by which the outcome and prior behavior can be connected is a defining feature of immanent justice reasoning (for an overview, see Callan et al., 2014). In two studies, Callan et al. (2006) showed that people resort to immanent justice reasoning for both positive and negative outcomes. Specifically, in their first study, they presented participants with a scenario in which a man named David was seriously injured in a car accident. Half of the participants learned that David was having an extramarital affair with a travel agent. The other half of the participants was told that David was on his way to a travel agent to plan a holiday with his family instead. They subsequently measured perceptions of a causal link between David's behavior and the car accident and found that people resort more to immanent justice reasoning when David had an extramarital affair than when he did not. They conceptually replicated these findings in a second study, in which they measured immanent justice for an undeserved positive outcome. Hence, immanent justice reasoning seems to be an additional coping strategy to deal with a threat to one's just-world belief.

In a more recent study, Harvey and Callan (2014) extended these findings by also including ultimate justice reasoning as a possible defensive strategy in the face of just-world threats. Ultimate justice reasoning is different from immanent justice reasoning in the fact that the former is focused forward, stressing that current undeserved outcomes will ultimately lead to a more meaningful life, while the latter is backward-looking, focusing on prior behaviors to explain current undeserved outcomes. Ultimate and immanent justice reasoning were found to be negatively correlated. Moreover, using a similar setup as Callan et al. (2006), they showed that people were more likely to resort to immanent justice reasoning when the victim was a "bad" person (such as a person who cheated) and to ultimate justice reasoning when the victim was a "good" person (such as a person who planned a holiday for his family). These effects were mediated by perceptions of deservingness.

Related to these two balancing strategies to preserve the just-world belief, Kay and colleagues put forward a related idea, which they termed compensatory rationalizations (for an overview, see Kay et al., 2007). Compensatory rationalizations are used to find a balance between positive and negative outcomes or traits. Specifically, Kay and Jost (2003) showed that people judge the system as more fair when they have been exposed to complementary stereotypes as opposed to non-complementary stereotypes. That is, with a short scenario they introduced a person to the participant. In this scenario they varied wealth and happiness of the person, such that he was either rich and unhappy, poor and happy, rich and happy, or poor and unhappy. The first two conditions constituted complementary stereotypes as they confer that no person can have it all. The latter two conditions constituted noncomplementary stereotypes. Reading the complementary scenarios led to higher ratings of the system as fair than reading the noncomplementary scenarios. Moreover, in a subsequent study Kay and Jost (2003) showed that noncomplementary scenarios led to an implicit concern with justice.

These two strategies of immanent and ultimate justice reasoning, on the one hand, and compensatory rationalizations, on the other, seem to contradict each other, leading to directly opposite predictions regarding people's reactions following victimizations (encompassing both individual cases and discrimination of groups in society). However, the two can be reconciled by specifying the conditions under which either will be adopted. Several studies (e.g., Gaucher et al., 2010; Kay, Jost, & Young, 2005) have looked into these factors and revealed, for instance, that while people resort to immanent justice reasoning or victim blaming mostly when the traits are causally relevant for the outcome (e.g., "poor and lazy"), while they resort more to compensatory rationalizations for traits that are irrelevant for the outcome (e.g., "poor, but happy", see for instance, Kay et al., 2005).

These studies show that following confrontations with injustice, people can adopt a broad range of strategies to uphold their faith in a just world. While the first studies on the justice motive focused mostly on victim blaming and derogation, more recent work began to uncover many other possible reactions to resolve a just-world threat, such as helping, immanent and ultimate justice reasoning, and compensatory rationalizations. Importantly, these reactions not only cover reactions to individual cases of injustice, but also include reactions to groups of people who are less well-off and reactions to inequality as we also saw in the studies on compensatory rationalizations. Hence, the role of justice beliefs in discrimination became the focus of research as well (see for instance work on system justification theory; Jost & Banaji, 1994; Jost, Banaji, & Nosek, 2004).

Going Beyond Deservingness

Within the justice motive literature, injustice is defined as deservingness. That is, the central tenet of just-world theory is that people have a fundamental need to believe that the world is a just place, which is defined as a world in which people get what they deserve. As such, it most closely aligns with a notion of equity, which can be defined as proportionality between an individual's outcome and his or her input, as discussed in the beginning of this chapter. However, several different justice principles can be distinguished (Deutsch, 1975). Specifically, Deutsch distinguished three principles of justice; equity, equality, and need. People can dispositionally as well as situationally differ in the principle that they apply.

Looking at the way in which people define justice, we also see such a differentiation. Studying over 5000 instances of injustice, provided by ordinary people, Finkel (2000) found that most people, when probed for instances of unfairness, refer to situations where innocence was punished, hard work was not rewarded, or an unfair advantage was given. However, people also referred to situations of unequal treatment as an instance of injustice. Hence, while many referred to instances related to equity and deservingness, other principles could and did play a role.

Importantly, when probed for instances of unfairness, people come up with instances of personally experienced unfairness but also with instances of observed unfairness, where the situation did not directly involve them. These impersonal situations encompassed more than half of the situations provided in Finkel's (2000) study and also seemed to increase with age, with children mentioning unfairness for others about half of the time, but young adults and elderly adults mentioning unfairness for others about two third of the time. These findings are especially interesting as they point to the idea that justice is not self-interested (a substantial amount of the time).

In line with people's commonsense notion of justice, Deutsch (1975) notes that equity is an economically oriented view of justice in which the rules of justice are met when an individual's outcome or reward is proportional to his or her input or contribution. Many (Western) societies do have such an economic orientation. Hence, in many societies the deservingness principle applies and people will live by the rules of equity (Martin, 1999).

However, in certain situations the principles of equality or need may be applied. Specifically, equality may be applied in solidarity-oriented groups or contexts and need in caring-oriented groups or contexts. One could easily imagine that while a person might adhere to an equity principle of justice in general, contexts do exist in which (s)he takes more of a caring or solidarity orientation, for instance in schools, in a home for the elderly, or when people have been struck by a natural disaster. In these instances, we are able to let go of our general justice principle of equity and focus more on the other person's needs. We want to teach our children, enhance the quality of life for the elderly and come to the aid of the persons who lost their homes due to a typhoon. These additional principles of justice deserve attention in future studies and should be incorporated in the justice motive. Focusing on when people adopt these principles of justice may also result in additional strategies in which people try to preserve their belief in a just world.

Concluding Remarks

In this chapter we have discussed how lay beliefs about social justice shape reactions following unjust events and more specifically reactions toward innocent victims of injustice. We have explained why people sometimes react in derogatory manners toward innocent victims, blaming them for their ill fates. We have discussed some important processes that shape these reactions and influence the construal and processing of these events. Moreover, we have discussed alternative reactions of helping and balancing strategies (such as immanent justice reasoning and compensatory rationalizations) that seem to be less detrimental for the victims involved. We finished this chapter with a discussion of varied perspectives of justice that lay people can adopt and that may be incorporated into the justice motive literature.

We chose to focus on work pertaining to people's justice motive and reactions toward victims specifically, as we believe this to be an important area of research in which justice judgments play an important social role and can have far-reaching consequences. It is important to note, however, that in choosing to do so, we did not provide a complete overview of the possible ways in which justice plays a role in the lives of people. For instance, we have not discussed reactions to personal encounters with unjust situations. While these reactions may to some degree overlap with the reactions discussed in this chapter, we only briefly touched upon related fields of study, for instance on distributive and procedural justice. A discussion of all research conducted on social justice was beyond the scope of this chapter, but we hope to have provided an overview of the array of ways in which lay people can react to innocent suffering and innocent victims specifically.

References

- Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). New York: Academic Press. doi:[10.1016/S0065-2601\(08\)60108-2](https://doi.org/10.1016/S0065-2601(08)60108-2)
- Bal, M. (2014). *Making sense of injustice: Benign and derogatory reactions to innocent victims*. Doctoral dissertation, Utrecht University.
- Bal, M., & Van den Bos, K. (2010). The role of perpetrator similarity in reactions toward innocent victims. *European Journal of Social Psychology, 40*, 957–969. doi:[10.1002/ejsp.668](https://doi.org/10.1002/ejsp.668)
- Bal, M., & Van den Bos, K. (2012). Blaming for a better future: Future orientation and associated intolerance of personal uncertainty lead to harsher reactions toward innocent victims. *Personality and Social Psychology Bulletin, 38*, 835–844. doi:[10.1177/0146167212442970](https://doi.org/10.1177/0146167212442970)
- Bal, M., & van den Bos, K. (2015). Putting the “I” and “Us” in justice: Derogatory and benevolent reactions toward innocent victims in self-focused and other-focused individuals. *Social Justice Research, 28*, 274–292. doi:[10.1007/s11211-015-0249-3](https://doi.org/10.1007/s11211-015-0249-3)
- Bal, M., & Van den Bos, K. (2016). *On human motivation and coping with threats to one's worldview: Transforming victim blaming into helping of innocent victims*. Manuscript submitted for publication.
- Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 282–316). Boston, MA: McGraw-Hill.
- Beauchamp, T. L. (2001). *Philosophical ethics: An introduction to moral philosophy* (3rd ed.). Boston: McGraw-Hill.
- Bègue, L., Charoillaux, M., Cochet, J., Cury, C., & De Suremain, F. (2008). Altruistic behavior and the bidimensional just world belief. *The American Journal of Psychology, 121*, 47–56. doi:[10.2307/20445443](https://doi.org/10.2307/20445443)
- Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.
- Callan, M. J., Ellard, J. H., & Nicol, J. E. (2006). The belief in a just world and immanent justice reasoning in adults. *Personality and Social Psychology Bulletin, 32*, 1646–1658. doi:[10.1177/0146167206292236](https://doi.org/10.1177/0146167206292236)
- Callan, M. J., Sutton, R. M., Harvey, A. J., & Dawtry, R. J. (2014). Immanent justice reasoning: Theory, research, and current directions. *Advances in Experimental Social Psychology, 49*, 105–161. doi:[10.1016/B978-0-12-800052-6.00002-0](https://doi.org/10.1016/B978-0-12-800052-6.00002-0)

- Chaiken, A. L., & Darley, J. M. (1973). Victim or perpetrator: Defensive attribution of responsibility and the need for order and justice. *Journal of Personality and Social Psychology*, 25, 268–275. doi:[10.1037/h0033948](https://doi.org/10.1037/h0033948)
- 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. doi:[10.1177/0146167299025002007](https://doi.org/10.1177/0146167299025002007)
- Cohen, R. L. (Ed.). (1986). *Justice: Views from the social sciences*. New York: Plenum.
- Cohen, R. L. (2016). Restorative justice. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 257–272). New York: Springer.
- Correia, I., Vala, J., & Aguiar, P. (2001). The effects of belief in a just world and victim's innocence on secondary victimization, judgements of justice and deservingness. *Social Justice Research*, 14, 327–342. doi:[10.1023/A:1014324125095](https://doi.org/10.1023/A:1014324125095)
- Correia, I., Vala, J., & Aguiar, P. (2007). Victim's innocence, social categorization, and the threat to the belief in a just world. *Journal of Experimental Social Psychology*, 43, 31–38. doi:[10.1016/j.jesp.2005.12.010](https://doi.org/10.1016/j.jesp.2005.12.010)
- DePalma, M. T., Madey, S. F., Tillman, T. C., & Wheeler, J. (1999). Perceived patient responsibility and belief in a just world affect helping. *Basic and Applied Social Psychology*, 21, 131–137. doi:[10.1207/S15324834BA210205](https://doi.org/10.1207/S15324834BA210205)
- Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, 31, 137–149. doi:[10.1111/j.1540-4560.1975.tb01000.x](https://doi.org/10.1111/j.1540-4560.1975.tb01000.x)
- Finkel, N. J. (2000). But it's not fair! Commonsense notions of unfairness. *Psychology, Public Policy, and Law*, 7, 898–952. doi:[10.1037/1076-8971.6.4.898](https://doi.org/10.1037/1076-8971.6.4.898)
- Folger, R. (Ed.). (1984). *The sense of injustice: Social psychological perspectives*. New York: Plenum.
- Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences*, 34, 795–817. doi:[10.1016/S0191-8869\(02\)00072-7](https://doi.org/10.1016/S0191-8869(02)00072-7)
- Gaucher, D., Hafer, C. L., Kay, A. C., & Davidenko, N. (2010). Compensatory rationalizations and the resolution of everyday undeserved outcomes. *Personality and Social Psychology Bulletin*, 36, 109–118. doi:[10.1177/0146167209351701](https://doi.org/10.1177/0146167209351701)
- Hafer, C. L. (2000a). Investment in long-term goals and commitment to just means drive the need to believe in a just world. *Personality and Social Psychology Bulletin*, 26, 1059–1073. doi:[10.1177/01461672002611004](https://doi.org/10.1177/01461672002611004)
- Hafer, C. L. (2000b). Do innocent victims threaten the belief in a just world? Evidence from a modified Stroop task. *Journal of Personality and Social Psychology*, 79, 165–173. doi:[10.1037/0022-3514.79.2.165](https://doi.org/10.1037/0022-3514.79.2.165)
- Hafer, C. L., & Bègue, L. (2005). Experimental research on just-world theory: Problems, developments, and future challenges. *Psychological Bulletin*, 131, 128–167. doi:[10.1037/0033-2909.131.1.128](https://doi.org/10.1037/0033-2909.131.1.128)
- Hafer, C. L., & Gosse, L. (2011). Predicting alternative strategies for preserving a belief in a just world: The case of repressive coping style. *European Journal of Social Psychology*, 41(6), 730–739. doi:[10.1002/ejsp.807](https://doi.org/10.1002/ejsp.807)
- Hafer, C. L., & Rubel, A. N. (2015). The why and how of defending belief in a just world. *Advances in Experimental Social Psychology*, 51, 41–96. doi:[10.1016/bs.aesp.2014.09.001](https://doi.org/10.1016/bs.aesp.2014.09.001)
- Hafer, C. L., & Sutton, R. (2016). Belief in a just world. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 145–160). New York: Springer.
- Harvey, A. J., Callan, M. J., & Matthews, W. J. (2014). How much does effortful thinking underlie observers' reactions to victimization? *Social Justice Research*, 27, 175–208. doi:[10.1007/s11211-014-0209-3](https://doi.org/10.1007/s11211-014-0209-3)
- Harvey, A. J., & Callan, M. J. (2014). Getting “just deserts” or seeing the “silver lining”: The relation between judgments of immanent and ultimate justice. *PlosOne*, 9(7), 1–8. doi:[10.1371/journal.pone.0101803](https://doi.org/10.1371/journal.pone.0101803)

- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33*, 1–27. doi:10.1111/j.2044-8309.1994.tb01008.x
- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25*, 881–919. doi:10.1111/j.1467-9221.2004.00402.x
- Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology, 88*, 498–509. doi:10.1037/0022-3514.88.3.498
- Kay, A. C., & Jost, J. T. (2003). Complementary justice: Effects of “poor but happy” and “poor” but honest” stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology, 85*, 823–837. doi:10.1037/0022-3514.85.5.823
- Kay, A. C., Jost, J. T., Mandisodza, A. N., Sherman, S. J., Petrocelli, J. V., & Johnson, A. L. (2007). Panglossian ideology in the service of system justification: How complementary stereotypes help us to rationalize inequality. *Advances in Experimental Social Psychology, 39*, 305–358. doi:10.1016/S0065-2601(06)39006-5
- Kay, A. C., Jost, J. T., & Young, S. (2005). Victim-derogation and victim-enhancement as alternate routes to system-justification. *Psychological Science, 16*, 240–246. doi:10.1111/j.0956-7976.2005.00810.x
- King, M. L., Jr. (1963). *Letter from Birmingham jail*. Retrieved from <https://kinginstitute.stanford.edu/king-papers/documents/letter-birmingham-jail>
- Kogut, T. (2011). Someone to blame: When identifying a victim decreases helping. *Journal of Experimental Social Psychology, 47*, 748–755. doi:10.1016/j.jesp.2011.02.011
- Laurin, K., Fitzsimons, G. M., & Kay, A. C. (2011). Social disadvantage and the self-regulatory function of justice beliefs. *Journal of Personality and Social Psychology, 149*–171. doi:10.1037/a0021343
- Lerner, M. J. (1977). The justice motive: Some hypotheses as to its origins and forms. *Journal of Personality, 45*, 1–52. doi:10.1111/j.1467-6494.1977.tb00591.x
- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Lerner, M. J., & Clayton, S. (2011). *Justice and self-interest: Two fundamental motives*. Cambridge: Cambridge University Press.
- Lerner, M. J., & Goldberg, J. H. (1999). When do decent people blame victims? The differing effects of the explicit-rational and implicit-experiential cognitive systems. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 627–640). New York: Guilford Press.
- Lerner, M. J., Miller, D. T., & Holmes, J. G. (1976). Deserving and the emergence of forms of justice. In L. Berkowitz & E. Walster (Eds.), *Advances in experimental social psychology* (Vol. 9, pp. 133–162). New York: Academic Press. doi:10.1016/S0065-2601(08)60060-X
- Lerner, M. J., & Simmons, C. H. (1966). Observer’s reaction to the “innocent victim”: Compassion or rejection? *Journal of Personality and Social Psychology, 4*, 203–210. doi:10.1037/h0023562
- Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum.
- Lipkus, I. M., Dalbert, C., & Siegler, I. C. (1996). The importance of distinguishing the belief in a just world for self versus for others: Implications for psychological well-being. *Personality and Social Psychology Bulletin, 22*, 666–677. doi:10.1177/0146167296227002
- Loseman, A., & Van den Bos, K. (2012). A self-regulation hypothesis of coping with an unjust world: Ego-depletion and self-affirmation as underlying aspects of blaming of innocent victims. *Social Justice Research, 25*, 1–13. doi:10.1007/s11211-012-0152-0
- Maas, M., & Van Den Bos, K. (2009). An affective-experiential perspective on reactions to fair and unfair events: Individual differences in affect intensity moderated by experiential mindsets. *Journal of Experimental Social Psychology, 45*(4), 667–675. doi:10.1016/j.jesp.2009.02.014

- Martin, L. L. (1999). I-D compensation theory: Some implications of trying to satisfy immediate-return needs in a delayed-return culture. *Psychological Inquiry*, *10*, 195–208. doi:[10.1207/S15327965PLI1003_1](https://doi.org/10.1207/S15327965PLI1003_1)
- Montada, L., & Schneider, A. (1989). Justice and emotional reactions to the disadvantaged. *Social Justice Research*, *3*, 313–344. doi:[10.1007/BF01048081](https://doi.org/10.1007/BF01048081)
- Novak, D. W., & Lerner, M. J. (1968). Rejection as a consequence of perceived similarity. *Journal of Personality and Social Psychology*, *9*, 147–152. doi:[10.1037/h0025850](https://doi.org/10.1037/h0025850)
- Opatow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*, 1–20. doi:[10.1111/j.1540-4560.1990.tb00268.x](https://doi.org/10.1111/j.1540-4560.1990.tb00268.x)
- Peters, S. L. (2006). *The social psychology of being better off than others*. Doctoral dissertation, Utrecht University.
- Rawls, J. (1971). *A theory of justice*. Oxford: Oxford University Press.
- Sabbagh, C., & Schmitt, M. (2016). Past, present, and future of social justice theory and research. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 1–11). New York: Springer.
- Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The justice sensitivity inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research*, *23*, 211–238. doi:[10.1007/s11211-010-0115-2](https://doi.org/10.1007/s11211-010-0115-2)
- Shaver, K. G. (1970). Defensive attribution: Effects of severity and relevance on the responsibility assigned for an accident. *Journal of Personality and Social Psychology*, *14*, 101–113. doi:[10.1037/h0028777](https://doi.org/10.1037/h0028777)
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*, 580–591. doi:[10.1177/0146167294205014](https://doi.org/10.1177/0146167294205014)
- Skitka, L. J., Aramovich, N., Lytle, B. L., & Sargis, E. (2009). Knitting together an elephant: An integrative approach to understanding the psychology of justice reasoning. In D. R. Bobocel, A. C. Kay, M. P. Zanna, & J. M. Olson (Eds.), *The psychology of justice and legitimacy: The Ontario symposium* (Vol. 11, pp. 1–26). Philadelphia, PA: Psychology Press.
- Stel, M., Van den Bos, K., & Bal, M. (2012). On mimicry and the psychology of the belief in a just world: Imitating the behaviors of others reduces the blaming of innocent victims. *Social Justice Research*, *25*, 14–24. doi:[10.1007/s11211-012-0150-2](https://doi.org/10.1007/s11211-012-0150-2)
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247. doi:[10.1207/s15327957pspr0803_1](https://doi.org/10.1207/s15327957pspr0803_1)
- Sutton, R. M., & Douglas, K. M. (2005). Justice for all, or just for me? More evidence of the importance of the self-other distinction in just-world beliefs. *Personality and Individual Differences*, *39*, 637–645. doi:[10.1016/j.paid.2005.02.010](https://doi.org/10.1016/j.paid.2005.02.010)
- Thomas, W. R., & Thomas, E. S. (1928). *The child in America*. New York: Knopf.
- Tyler, T. R. (1987). Conditions leading to value-expressive effects in judgments of procedural justice: A test of four models. *Journal of Personality and Social Psychology*, *52*, 333–344. doi:[10.1037/0022-3514.52.2.333](https://doi.org/10.1037/0022-3514.52.2.333)
- Tyler, T. R. (1989). The psychology of procedural justice: A test of the group-value model. *Journal of Personality and Social Psychology*, *57*, 830–838. doi:[10.1037/0022-3514.57.5.830](https://doi.org/10.1037/0022-3514.57.5.830)
- Tyler, T. R. (2006). *Why people obey the law*. Princeton, NJ: Princeton University Press.
- Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 115–191). San Diego, CA: Academic Press. doi:[10.1016/S0065-2601\(08\)60283-X](https://doi.org/10.1016/S0065-2601(08)60283-X)
- Van den Bos, K. (2005). What is responsible for the fair process effect? In J. Greenberg & J. A. Colquitt (Eds.), *Handbook of organizational justice: Fundamental questions about fairness in the workplace* (pp. 273–300). Mahwah, NJ: Erlbaum.
- Van den Bos, K. (2015). Humans making sense of alarming conditions: Psychological insight into the fair process effect. In R. S. Cropanzano & M. L. Ambrose (Eds.), *Oxford handbook of justice in work organizations* (pp. 403–417). New York: Oxford University Press.
- Van den Bos, K. (2007). Hot cognition and social justice judgments: The combined influence of cognitive and affective factors on the justice judgment process. In D. de Cremer (Ed.),

- Advances in the psychology of justice and affect* (pp. 59–82). Greenwich, CT: Information Age Publishing.
- Van den Bos, K., Ham, J., Lind, E. A., Simonis, M., Van Essen, W. J., & Rijkema, M. (2008). Justice and the human alarm system: The impact of exclamation points and flashing lights on the justice judgment process. *Journal of Experimental Social Psychology, 44*, 201–219. doi:[10.1016/j.jesp.2007.03.001](https://doi.org/10.1016/j.jesp.2007.03.001)
- Van den Bos, K. (2009). Making sense of life: The existential self trying to deal with personal uncertainty. *Psychological Inquiry, 20*, 197–217. doi:[10.1080/10478400903333411](https://doi.org/10.1080/10478400903333411)
- Van den Bos, K., & Maas, M. (2009). An affective-experiential perspective on reactions to fair and unfair events: Individual differences in affect intensity moderated by experiential mindsets. *Journal of Experimental Social Psychology, 45*, 667–675. doi:[10.1016/j.jesp.2009.02.014](https://doi.org/10.1016/j.jesp.2009.02.014)
- Van Prooijen, J. W. (2013). Individualistic and social motives for justice judgments. *Annals of the New York Academy of Sciences, 1299*, 60–67. doi:[10.1111/nyas.12143](https://doi.org/10.1111/nyas.12143)
- Van Prooijen, J. W., & Van den Bos, K. (2009). We blame innocent victims more than I do: Self-construal level moderated responses to just-world threats. *Personality and Social Psychology Bulletin, 35*, 1528–1539. doi:[10.1177/0146167209344728](https://doi.org/10.1177/0146167209344728)
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology, 3*, 73–79. doi:[10.1037/h0022733](https://doi.org/10.1037/h0022733)
- Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*. Boston: Allyn & Bacon.
- Wenzel, M., & Mummendey, A. (1996). Positive-negative asymmetry of social discrimination: A normative analysis of differential evaluations of in-group and out-group on positive and negative attributes. *British Journal of Social Psychology, 35*, 493–507. doi:[10.1111/j.2044-8309.1996.tb01110.x](https://doi.org/10.1111/j.2044-8309.1996.tb01110.x)
- Wenzel, M., & Okimoto, T. G. (2016). Retributive justice. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 237–256). New York: Springer.

Part III
Insights into Lay Theories About the
Metaphysical or Supernatural

Antecedents, Manifestations, and Consequences of Belief in Mind–Body Dualism

Matthias Forstmann and Pascal Burgmer

Just like professional scientists, laypeople naturally generate hypotheses about the world and gather data through observation (e.g., Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1992). Yet, unlike scientists, who ideally deduce their hypotheses from formal theories, laypeople typically derive theirs from cumulatively acquired belief systems. That is, people use the information they acquire throughout their lives to form various lay theories about the world they live in, and use these belief systems to interpret and organize novel information (e.g., Kelley, 1967; Sloman, 2009). People have been found to hold common-sense beliefs and lay theories about a great number of things they encounter in various spheres of life, and they rely on these beliefs to guide their behaviors across a wide range of situations (Molden & Dweck, 2006). For example, people have elaborate theories about whether or not the world is a just place, where people get what they deserve (e.g., Callan, Kay, Davidenko, & Ellard, 2009; Lerner, 1980; Bal & van den Bos, this volume), whether personality traits are malleable or fixed (e.g., Dweck & Leggett, 1988; Molden & Dweck, 2006; Burnette, Hoyt, & Orvidas, this volume; Rattan & Georgeac, this volume), or whether willpower is a limited resource that can be fatigued (Job, this volume; Job, Dweck, & Walton, 2010).

Sometimes these beliefs can pertain to rather philosophical topics that do not offer clear-cut answers: do we have free will, or are our actions determined? Is our mind simply what our brain does, or is there an immaterial self that exists

M. Forstmann (✉)

Department of Psychology, Yale University, 2 Hillhouse Ave,
New Haven, CT 06520-8205, USA
e-mail: matthias.forstmann@yale.edu

P. Burgmer

Department of Psychology, University of Cologne, Cologne, Germany

independently of our body? Recently, scholars from various fields have taken an interdisciplinary approach to understanding how lay people think about these questions, commonly referred to as experimental philosophy (Knobe, 2007): social psychologists, cognitive, and developmental scientists, and experimental-minded philosophers have started collaborative work to investigate the antecedents and consequences of lay theories about metaphysical issues, such as free will, morality, or intentionality. One particular domain that has become a focus of attention of experimental philosophers is the *philosophy of mind*. It concerns questions about what the mind is, why we have conscious experience, and—if it even exists—how an immaterial mind may relate to the human body.

Terminology

In contrast to some of the other chapters in this volume, we chose to speak of *beliefs* rather than lay theories, naïve theories, implicit theories, or mindsets. Although these terms are oftentimes used interchangeably, we consider belief to be the most adequate term to describe the variety of phenomena discussed in the literature. Regardless of whether or not they are justified or true (Starmans & Friedman, 2012), beliefs can generally be considered sets of assumptions that a person has about the current state of the world or about certain rules underlying reality. While this definition is undoubtedly rather broad, it allows to cover a host of assumptions people may have, ranging from metaphysical propositions regarding the existence of free will to mundane assumptions about the current location of their car keys. Contrary to that, the term lay theory carries with it a certain scientific connotation, based on the idea of the lay person as a “naïve” scientist. This term is in our opinion more suited for specific, complex belief systems—based on testable hypotheses—that aim for a non-metaphysical, rational, abstract understanding of the social world. A theory technically refers to a sophisticated explanatory framework, which does not necessarily apply to some of the beliefs we are talking about in this chapter. Further, whether these various beliefs fulfill the criteria to be referred to as implicit (e.g., people’s inability to report their existence or operation; see Greenwald & Banaji, 1995) should be decided on a case-by-case basis (see Wegener & Petty, 1998).

Mind–Body Dualism

Although hinted at by Plato in his *Theory of Forms* (Plato, 360 B.C./1977), it was René Descartes who first extensively wrote about what later became known as the philosophical theory of mind–body dualism or the mind–body problem

(Descartes, 1641/1641). Simplified, Descartes argues that the nature of the mind (e.g., thinking, spatially and temporally unrestricted) is fundamentally different from the nature of the body (e.g., nonthinking, spatially, and temporally restricted). He wrote that he could clearly conceive of minds without bodies and bodies without minds. Therefore, he concludes, mind and body must be two conceptually different entities made up of two fundamentally different kinds of substance. Evidently, this reasoning, which had a profound impact on the philosophy of the eighteenth and nineteenth century, was entirely based on his own introspective, phenomenological experience. After all, Descartes merely describes (in undeniably well-conceived terms) what it *feels like* to have a mind, a feeling that all of us can intuitively relate to.

The experience of consciousness has a distinct quality, in philosophy referred to as *qualia*, and is inherently subjective. Although one could, for example, easily describe to a blind person what a color is in physical terms, it is impossible to communicate how perceiving a certain color feels like. As Thomas Nagel puts it, since the physical world deals in objectives and *qualia* is per definition fundamentally subjective, we may never know what it *feels like* to be a bat, no matter how much we observe or study it (Nagel, 1974). As a result, although the knowledge gathered by the natural sciences such as physics, biology, and chemistry allows us to understand many of the physical processes that are at work in the human +brain, the discourse on the nature of mind–body relations continues even today (e.g., Chalmers, 1995; Kim, 2000).

But philosophical debates notwithstanding, how do lay people perceive the relation between their and others' mind and body? Do they instinctively follow Descartes' reasoning or are they less convinced by the prospect of a mind that is separate from a body? Or do they have no opinion on this matter at all, unless they are directly questioned about it?

When speaking about beliefs in mind–body dualism, it seems conceptually reasonable to differentiate between two separate, yet most likely interrelated, constructs: *intuitive beliefs* in mind–body dualism (or *implicit dualism*, as put by Uhlmann, Poehlman, & Bargh, 2008), which are, (a) natural inclinations to understand others' minds as being non-contingent on their brains and (b) the self-perception of occupying rather than being one's physical body, and *explicit beliefs* in mind–body dualism, that is, deliberate assumptions about how mental states may be more than just the product of a physical body. The latter include, for example, philosophical considerations regarding this issue, (more indirectly) religious beliefs that posit the existence of immortal souls or an afterlife, or beliefs in the ability of minds to exist independently of a body, such as in the case of ghosts or spirit possession. Explicit beliefs are thus based on the reflective thinking: they are deliberate, acquired, stable, and rather rigid assumptions about mind–body relations. Intuitive beliefs, in contrast, are results of phenomenological experiences or "feelings": they are mostly unconscious, structurally innate, situationally variable, and therefore rather flexible perceptions of others and ourselves.

Intuitive (or Implicit) Beliefs in Mind–Body Dualism

Although people may differ with regard to their explicit endorsement of dualistic positions, it has been argued that all humans seem to be—for a lack of a better term—“natural-born dualists” (Bloom, 2004, p. xiii), and that the ability to differentiate mind and body can be considered a defining capacity of the human species (Povinelli & Bering, 2002; Suddendorf & Whiten, 2001). That is, at least on an intuitive level, we all seem to share an understanding of minds being somehow independent of physical bodies. We seem to naturally understand others’ minds to be not entirely contingent on their physical constitution, and have the same intuition with regard to our own mind and body, regardless of our explicit beliefs.

Consequently, it seems reasonable to differentiate between two types of intuitive mind–body dualism: self-oriented and other-oriented type (Fig. 1), which may or may not be causally related to one another. In the following, we will try to outline the manifestations and causes of these two seemingly universal intuitions.

Dualistic Views on Others’ Minds

When it comes to others’ minds, our dualistic intuitions become evident in rather mundane aspects of our daily lives, for instance in pop-cultural fiction that involves body transformation or anthropomorphism—that is, attributing humanlike properties, characteristics, or mental states to real or imagined nonhuman agents and

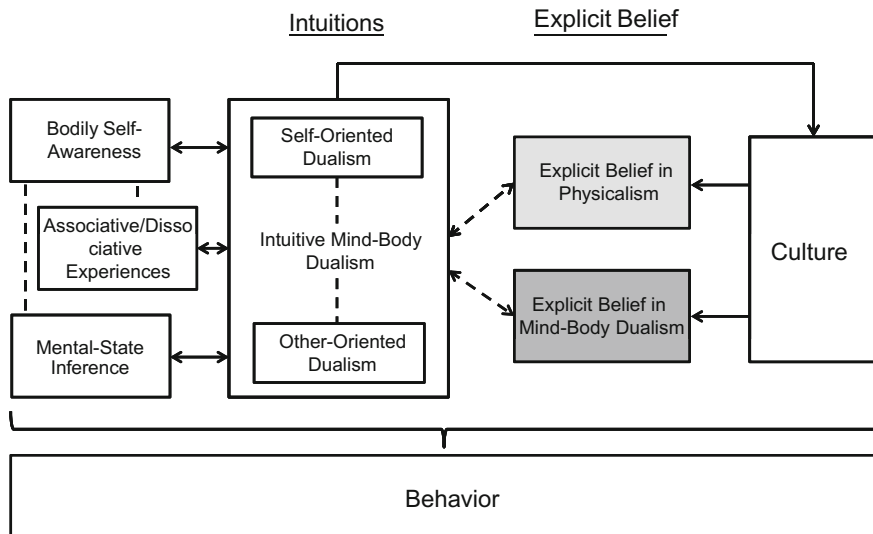


Fig. 1 Intuitive and explicit belief in mind–body dualism

objects (Epley, Waytz, & Cacioppo, 2007, p. 865). Although we know that they are not likely to actually become a reality, we effortlessly comprehend a parable in which a man wakes up inhabiting the body of an insect while retaining his former mental states (Kafka, 1915/1996), a scary movie about a murderous doll possessed by an evil spirit, or a TV commercial in which a sofa loudly proclaims the latest discounts of a furniture store. As different as these cases may be, they are all instances in which we intuitively infer the presence of mental states in entities that clearly lack (human) brains, as if their minds were somehow not contingent on their physical constitution. After all, how could a sofa even know what a discount is if it does not fulfill the necessary physiological requirements to “know” anything at all? Importantly, exposure to these innocent tales and cultural products does not puzzle or bewilder us. We do not find them confusing and they do not leave us wondering what on earth is actually going on (Bloom, 2004). For us, the mere fact that a being seems to have a mind that it could not (or should not) have seems to be no reason for concern.

Mind–Body Dualism as Cognitive Default

As stated above, the root of this effortless comprehension seems to be our natural inclination to intuitively perceive minds to not solely be functional products of specific brain states. In fact, recent work in developmental and social psychology strongly suggests that both adults and children intuitively apply this kind of unconscious dualistic reasoning when reflecting about other entities’ mental states, and that this intuitive dualism can in fact be considered a cognitive default—that is, a habitual way of thinking about this issue (e.g., Bering, 2006). In some of our own work (Forstmann & Burgmer, 2015), we explored this issue by presenting participants with various thought experiments, which are similar to the classic teleportation thought experiments introduced by Derek Parfit (1984) that were later adapted by developmental psychologists to study mind–body dualism in children (Hood & Bloom, 2008; Hood, Gjersoe, & Bloom, 2012).

Specifically, participants were told about a hypothetical future in which a novel duplication device was able to duplicate any kind of object in a matter of seconds by scanning it, atom for atom, and then using the information gathered to assemble a perfect duplicate at a second location from basic chemical elements. Emphasis was put on the fact that the machine was 100% accurate and reliable, in that it created a 100% identical physical copy of the original. Subsequently, participants were told about a lab hamster that was duplicated by the device. This hamster was described using twelve attributes, six of which were mental (e.g., the hamster is afraid of the lab intern; the hamster vividly remembers his hamster sister), and six of which were physical in nature (e.g., the hamster has a limp; the hamster has a complicated brain tumor). Participants were then asked to imagine that the duplication procedure was conducted, and were asked about how much they thought the twelve attributes (still) applied to both the original and duplicate hamster. Results indicated that while most adults considered the physical properties of the duplicate

hamster to have remained more or less unchanged (including the presence of a brain tumor), they considered it to have lost certain mental states during the procedure, such as memories and acquired knowledge structures. In other words, despite having an identical brain, participants ascribed to the duplicate hamster vastly different mental states as compared to the original. They thus revealed an intuitive inclination to differentiate mental states from physical properties, that is, a dualistic view on minds and bodies.

Importantly, this effect was more pronounced under conditions of heightened cognitive load. Cognitive load refers to the amount of cognitive resources that are being used by the working memory at any given time (Sweller, 1988). It is thus equatable with the level of mental effort a person exerts. As cognitive resources are limited, the more people are under cognitive load, the less resources are available for controlled mental tasks that need to be performed simultaneously (Sweller, 1988). Thus, by putting people under cognitive load—for example, by having them remember a complicated string of random letters while working on a task—one can lower the amount of available cognitive resources and thereby indirectly promote the use of automatic (or default) mental processes, which do not require these resources. In our study, participants who were put under cognitive load indicated greater levels of intuitive mind–body dualism than did participants whose resources had not been taxed. One way to interpret these results is that the intuitive mind–body dualism that most participants revealed can be considered a *cognitive default*, which we all share to some extent. The data suggest that we need cognitive resources in order to deliberately override these natural intuitions. Once these resources are strained, intuitive dualism increases.

Similarly supporting this notion, intuitive mind–body dualism increased when participants were procedurally primed with an intuitive thinking style prior to working on the thought experiment. That is, participants were experimentally manipulated to rely more on their (automatic) intuitions—or their (resource-demanding) analytic thinking, respectively—when responding to the task, by having them recall an instance in the past in which they relied on their intuition to solve a problem. Recall of such a thinking style is believed to carry over to subsequent tasks (Gervais & Norenzayan, 2012). Again, participants who relied on their intuitions revealed greater levels of intuitive mind–body dualism than did participants who relied on analytical thinking.

This cognitive default seems to manifest at rather early stages in human development. In a study by Bering and Bjorklund (2004), for example, children were told a story about a mouse that was eaten by a scrupulous alligator. Then, these children were asked about their opinion with regard to the continuing psychological functioning of the dead mouse. In line with a dualistic view on minds and bodies, the children were more likely to ascribe to the dead mouse continuing emotional, epistemic and desire states (e.g., the mouse can still love its mother) than (psycho)biological and perceptual states (e.g., the mouse cannot feel hunger anymore).

For children, the mere fact that the dead mouse does not have a brain is no reason to assume a lack of certain important mental states. In fact, it seems that even

rudimentary knowledge about the brain is not sufficient to override early dualistic intuitions. While 3–5 year-old children already understand that a brain is needed for thinking, planning, and remembering things, they tend to regard it as a tool that is utilized to perform these operations (Johnson & Wellman, 1982). Consequently, they still do not consider the brain a necessity for performing simple actions such as brushing one's teeth or walking down a street. Similarly, most children of that age do not believe that a brain is necessary for pretending to be kangaroo or a tree, but, for example, for deciding prior to the act on how to perform these imitations (Lillard, 1996).

While by kindergarten age, children know that individual body parts have certain functions, the knowledge that the brain itself is entirely responsible for perception, cognition, and behavior is only gradually acquired over the course of a few years (Johnson, 1990), presumably via exposure to culturally shared explicit knowledge about how the human body operates (Marshall & Comalli, 2012). This developmental process was, for instance, investigated using (fictitious) brain transplant studies, in which children of different age groups were asked about what they thought would happen if a brain of one animal was replaced with the brain of another animal (Gottfried, Gelman, & Schultz, 1999). Older children and adults understood that the brain “contained” thoughts and feelings, and that these states would thus transfer to the new host. 5-year olds, however, knew that having a brain is a requirement for performing mental tasks such as thinking and remembering, but believed that any brain would suffice to give rise to cognitions and memories that matched the specific animal's category membership. That is, these children had essentialist assumptions about the animals, in that they spontaneously made inferences about internal features and non-visible functions (such as mental states) purely based on the category membership (Gelman, 2004). As stated by one of the children, a horse that is equipped with a cow's brain thinks about running fast (as opposed to giving milk) because “it has a brain now” (Gottfried et al., 1999). In other words, younger children treated mental states to be something that cannot be equated with the brain, even after they learned that a brain is factually responsible for thinking and feeling. This further supports the idea that a dualistic view on mind and body is a cognitive default that is already present in young children, and that is only later replaced (or rather suppressed) by more elaborate explicit knowledge about the brain and its relation to mental states.

Mind–Body Dualism as a Consequence of Theory of Mind Development

But where does this intuitive conception of others' minds being independent of their brains or bodies come from? As some argue, the intuitive dualism that we all seem to share can be conceived of as a side effect of a collection of basic cognitive processes that are often collectively referred to as a *Theory of Mind* (Bering, 2006; Bloom, 2004). A Theory of Mind refers to our ability to understand that others have mental states—including emotions, knowledge, and memories—that can be incongruent with our own, as well as to our ability to make use of this insight by

applying it to our interpretations and predictions of others' behavior (Premack & Woodruff, 1978; Wellman, Cross, & Watson, 2001). While this ability undoubtedly constituted an important evolutionary advantage—after all, knowing what someone else is up to can be crucial for survival—it may also have inadvertently enabled the formation of dualistic beliefs (Bering, 2006, 2011; Povinelli & Bering, 2002).

More precisely, in order to be able to better predict what is happening in their surroundings, all humans possess some sort of pre-built-in agency-detection device (Guthrie, 1993), and readily ascribe intentions, goals, and desires to other human beings. These processes are such a fundamental part of our cognitive architecture that we even ascribe the same mental faculties to self-propelled, nonhuman entities, such as simple shapes on a computer screen (Gergely & Csibra, 2003) or moving triangles (Heider & Simmel, 1944). Even infants are already able to distinguish between intentional and nonintentional movement (Woodward, 1998, 1999), and are able to identify and imitate goal-relevant aspects of complete or even incomplete actions performed by human agents (Hamlin, Hallinan, & Woodward, 2008). At twelve months of age, children take what Dennett (1971) refers to as the *intentional stance*. They clearly represent an agent's goals, intentions, and desires, and can distinguish between rational and irrational means to reach these goals (Gergely, Nádasdy, Csibra, & Bíró, 1995).

In order to be able to do this, humans necessarily have to make inferences about mental processes that are inaccessible to them, and that they cannot directly perceive with their senses. That is, while others show observable behavior, we know that we can only speculate about the covert mental states (goals, intentions, etc.) that caused this behavior. While we see how others behave, and hear what they say, we can only assume what goes on in the enigmatic mind that must be in charge of these actions. Consequently, the development of a Theory of Mind may inadvertently promote the development of two different modes of construal: one exclusively dealing with the physical (i.e., observable), the other with the social (i.e., unobservable) world (Bloom, 2004). It may lead us to think of the mental and the physical as two separate realms, and thereby foster the development of intuitive beliefs in mind–body dualism.

This conceptual differentiation between social and nonsocial entities can already be observed in infants (e.g., Legerstee, 1992). In a study by Kuhlmeier, Bloom, and Wynn (2004), the authors employed an expectation-violation paradigm to investigate whether or not 5-month-old-infants perceive humans as material objects. To do so, they analyzed looking times of infants—a proxy for expectation-violation with longer looking times indicating greater confusion about the situation they just witnessed—after they saw either a physical object or a human being move in a manner not compatible with the laws of physics. Specifically, the infants were seated in front of a screen with three slits, through which the rest of the room could be observed. When an object that moved behind the screen initially appeared behind the first and then behind the third slit—skipping the second—the infants were clearly surprised, revealing a seemingly innate “naïve theory of physics” (Spelke, 1991). When a human being moved in this pattern (realized with the help of twin experimenters), the infants reacted less surprised. While they do not

constitute direct manifestations of dualistic belief, these results suggest that from early on in their lives, humans draw a clear distinction between animate and inanimate objects. That is, they use two modes of construal for physical objects with and without minds, with different rules applying to each. The mere fact that infants already make a mind/no-mind distinction reveals how fundamental this differentiation is for us, and how we are hardwired to think of the mental as somehow *different* from the physical world.

Thus, it seems that our natural tendency to differentiate others' unobservable minds from their observable bodies, a key element in mental state inference and an evolutionary necessity, lays the foundation for intuitive dualistic beliefs. It allows us to accept the aforementioned concepts of possessed dolls or talking sofas, as we treat mental life as something that is just not part of the physical world that we perceive with our senses. Supporting this notion, our own work suggests that both belief in mind–body dualism and mental state inference (that is, perceptual and conceptual perspective-taking), are two closely interrelated constructs (Burgmer, Forstmann, Todd, & Mussweiler, 2016). Specifically, participants whose explicit beliefs in mind–body dualism were experimentally strengthened (vs. weakened) by reading a short vignette text describing the philosophical position of mind–body dualism (vs. a text about materialistic monism) (cf. Forstmann, Burgmer, & Mussweiler, 2012) were more likely to spontaneously adopt another person's visual vantage point when determining the spatial location of an object. Specifically, participants were presented with a photograph of a person standing in front of a table, facing the camera, with a book lying on one side of the table. Participants answered various questions with regard to the content and properties of the photograph (e.g., quality of the photograph), among them a critical item asking on which side of the table the book was. This item can either be answered in self-oriented or in other-oriented fashion. Our results indicate that participants in the dualism condition were more likely to give other-oriented responses by adapting the visual perspective of the person in the picture, while participants in the physicalism condition tended to give self-oriented responses by indicating the location of the book from their own vantage point.

Further, participants who read about mind–body dualism were more likely to overcome their egocentric bias in a *false belief* reasoning task—a task that assesses people's appreciation that others may hold false beliefs about reality which deviate from their own. In this paradigm, participants are told a story about a girl named Vicki who puts her violin in one of four boxes before going out to play. In her absence, her sister moves the violin to a different box—a fact that Vicki is unaware of. Upon her return, she wants to continue playing her instrument. Participants are then asked to indicate for each of the four boxes how likely it is that Vicki will look for her instrument there first. Importantly, they need to suppress their own privileged knowledge (i.e., that the violin is now in a different location) in order to estimate that Vicki will first look for the instrument in the box in which she put the violin earlier. Overcoming such an egocentric “curse of knowledge” is an indicator of successfully taking other people's perspectives (Birch & Bloom, 2007). Our findings thus indicate that dualism helps people suppress their own privileged

knowledge about a situation when inferring the mental content of another individual (Birch & Bloom, 2007; Wimmer & Perner, 1983). In line with the argument outlined above, we further found that this effect can indeed partially be explained by participants' perceived necessity to infer mental states in others, which was heightened for people whose dualistic beliefs had been strengthened by our experimental manipulation. In other words, mind–body dualism increased participants' awareness about the fact that mental content of others needs to be inferred—that one needs to look beyond what can be seen—which ultimately increased their inclination to engage in perspective-taking.

Dualistic Views on Our Own Mind

Our intuitions about mind–body relations obviously do not solely concern the minds and bodies of others, but also our own. We also do not perceive ourselves to be a bundle of firing neurons, or an accumulation of electrical signals in a lump of protein. Rather, we feel as if our mind has a distinct quality (the aforementioned qualia) and occupies our physical body, like a vessel, and that we only use this vessel to navigate the material world (Bloom, 2004). We do not feel as if we are “machines made of meat” (Bloom, 2004), but rather as if our self resides at a certain location within our head. In fact, it seems as if lay people intuitively invoke the metaphor of the eyes as the window to the soul: when asked about which of various drawn objects was closest to a person displayed, children and adults tended to choose the object that was closest to the eyes of the person even though all objects were equally far away from *some* part of the person's body (Starmans & Bloom, 2012).

In our experience, we are intentional agents possessing free will, and we perceive our mind to be the source of this freedom. As Preston, Grey, and Wegner (2006) put it, we have a “compelling feeling of personal causation that accompanies almost every action we take, and suggests that an immaterial self is in charge of the physical body” (p. 482). The phenomenological experience of free will poses a stark contrast to the physicalistic view on minds and bodies that only truly allows for a deterministic worldview (Nahmias, Shepard, & Reuter, 2014). If everything we experience is completely explained by neurochemistry, where could free will come into play?

And not only do we feel free and in charge of what our physical body does at any given point in time, our mind also feels like it extends beyond the here and now. It enables us to imagine counterfactual scenarios, fantasize about far-away places, dream up impossible objects, or remember better times (e.g., Smallwood & Schooler, 2006). In these instances, we have the phenomenological experience of “seeing” and “hearing” things in our head that are not the result of any sensory input we receive from the outside. Similar to Descartes' (1641/1984) description, our mind feels temporally and spatially unrestricted, and our mental life feels fundamentally different from our immediate perception.

As with our intuition about others' minds and bodies, this subjective experience of mind–body dualism seems likewise to be rooted in our basic cognitive architecture. When we develop a Theory of Mind, it not only means that we understand that others have minds that we need to infer to understand their view of the world, it also means that we know that our own mental states are similarly inaccessible to the outside world, thereby constituting an individual realm on their own.

Developmental research points toward the notion that children know from rather early on in their lives that their mental states are accessible only to themselves, for example when they pretend to be in pain (Antony, 2006), or when they intentionally deceive others (Hala, Chandler, & Fritz, 1991; Wimmer & Perner, 1983). Deception requires the insight that one's mind is private and that others are not aware of the actual knowledge one possesses. Thus, children's ability to use deception can be understood as a marker for their Theory of Mind development, as it reveals a strategic use of knowledge about others' diverging mental states (Hala et al., 1991). When it comes to their own minds, young children reveal an ontological distinction between their own subjective experience and the external world rather early in their development. For example, 4-year-old children know that dreams are fictional, that they are the only ones subjectively experiencing them, and that the content of their dreams is non-perceptible to others (Woolley & Wellman, 1992). Yet, they also know that objects imagined in waking hours are immaterial and similarly only perceptible to themselves (Wellman & Estes, 1986; Estes, 1994). They already understand that thought-about objects are not factually inside their head, whereas they know that swallowed objects are indeed resting inside their bodies (Watson, Gelman, & Wellman, 1998).

Wellman and Estes (1994) explored children's ontological distinction between their mind and the external world in more detail. According to the authors, physical objects and mentally represented objects primarily differ on four critical dimensions. The *behavioral-sensory dimension*, that is, whether an entity can be touched, seen, or acted upon; the *public-existence dimension*, that is, whether others can similarly perceive and act upon this entity; the *consistent-existence dimension*, that is, whether the entity persists over time; and the *realism dimension*, that is, whether the entity has to adhere to certain laws of physics. Wellman and Estes found that 3-year olds are able to clearly distinguish physically existing from imagined objects on all four dimensions. They understand that imagined objects cannot be touched, are only perceived by themselves, only exist while they think about them, and do not have to be realistic. Children of this age also engage in a rudimentary form of introspection, in that they reflect on and discuss mental imagery (Estes, 1994). Combined, these results indicate that a subjective phenomenological distinction between one's mind and body develops at rather early stages of human development. Such intuitions contribute to a perceived separateness of one's mind and body, thereby contributing to the formation of intuitive self-oriented dualistic beliefs.

Bodily Self-awareness and Dissociation

In addition to the ontological distinction between mental imagery and physical reality, an intuitive belief in mind–body dualism may critically depend upon one’s perceived relationship with one’s body, be it a feeling of dissociation or association. On the one hand, we frequently make certain experiences that foster a perceived dissociation from our body. As outlined above, in cases of dreaming or mind-wandering (Smallwood & Schooler, 2006), our phenomenological experience is dissociated from our sensory experience. Further, we experience instances in which we feel to be “in charge” of our body. We possess the ability to mentally suppress undesired psycho-physiological states such as pain, hunger, or fear (e.g., Gross, 1998), and are able to calm our body down, for instance by shifting our attention or by engaging in mental relaxation. On the other hand, we also make experiences that strengthen the perception of *being* our body rather than of just occupying it: for example, we may experience immediate effects of ingesting drugs or medication on our mental life (e.g., DeWall et al., 2010). Drinking a bottle of wine, technically a strictly physical act involving the mere locomotion of a liquid, has almost immediate effects on our subjective phenomenological experience of the world. Many of us also notice when bodily states such as hunger or tiredness alter our mood in a negative manner (e.g., Stucke & Baumeister, 2006), when sleep deprivation inhibits our performance (Pilcher & Huffcutt, 1996), or when bodily fatigue affects our self-regulation abilities (Nilsson et al., 2005). These are all cases in which the association between body and mind is made especially salient. It is well conceivable that this salience may situationally counter our default dualistic intuitions.

In other words, intuitive belief in self-oriented mind–body dualism is likely to be heavily affected by situational factors related to bodily self-awareness—or the corresponding salience of the connection between our mind and our body. Bodily self-awareness is theorized to primarily comprise two subcomponents: *body ownership*, that is, the experience of owning a body, and *self-location*, that is, the experience of being a body with a given location in space and time (Serino et al., 2013). Especially, the latter seems to be crucial for our intuitions about mind–body relations. This aspect of self-awareness is tightly linked to the perception of one’s own bodily states: experiments show that people’s perceived self-location can be shifted toward the location of a virtual body by disintegrating somatosensory (i.e., proprioceptive and tactile) and visual stimulation (Ionta et al., 2011; Lenggenhager, Tadi, Metzinger, & Blanke, 2007). Put differently, we seem to use our own bodily states in combination with visual cues as information about where our self is located and which physical entities (such as body parts) we include in our self. For example, by simultaneously applying haptic stimulation to real and artificial limbs (“rubber hand illusion”; see Botvinick & Cohen, 1998), it is possible to induce the perception of having a third arm (Guterstam, Petkova, & Ehrsson, 2011). Additionally, by artificially altering participants’ visual perception in a way that makes them look at their own bodies from an observer perspective, it is even

possible to induce out-of-body experiences, that is, states of complete dissociation from any form of physical body (Ehrsson, 2007). Therefore, it is reasonable to assume that certain associative and dissociative experiences we make on a daily basis may strengthen or weaken the degree to which we incorporate our body in our self.

The phenomenon of self-location can be neurologically mapped to a brain region referred to as the temporal–parietal junction (Serino et al., 2013), a region that (among others) was similarly found to play a key role in perspective-taking and Theory of Mind (e.g., Saxe & Kanwisher, 2003). Cortical activation in these areas was further linked to dissociative states (e.g., Simeon, Guralnik, Schmeidler, Sirof, & Knutelska, 2001) and the aforementioned out-of-body experiences, that is, disturbed own-body perceptions (Blanke & Arzy, 2005; Pavani, Spence, & Driver, 2000). In fact, a study found that participants who reported frequent out-of-body experiences also performed better in a mental rotation task that required taking an exocentric perspective than did participants who did not report such experiences (Blackmore, 1987). This finding illustrates once more the intimate connection between certain dualistic experiences and perspective-taking. In conclusion, mental state inference, bodily self-perception, and therefore intuitive mind–body dualism seem to share certain neurological underpinnings. It seems as if the same brain regions responsible for reasoning about other people’s minds (which involves dissociating oneself from one’s own perspective) are involved in dissociating one’s own mind from one’s body, once more hinting at a close relationship between intuitive mind–body dualism and mental state inference.

To conclude, certain fundamental cognitive and perceptual processes that all humans are naturally equipped with are likely to contribute to both self- and other-oriented intuitive mind–body dualism. These processes—most likely associated with cortical activation in the temporal parietal junction in the human brain—concern perspective-taking or Theory of Mind on the one side, and bodily self-awareness or self-localization the other side (see Fig. 1).

Explicit Beliefs in Mind–Body Dualism

Considering that our intuitions about minds and bodies are assumed to be a function of basic human cognitive and perceptual processes (i.e., a Theory of Mind, bodily self-awareness) and common human experiences (e.g., mind-wandering, day-dreaming), it is no wonder that *explicit* belief systems based on these intuitions can be found in virtually all human cultures throughout history (e.g., Chudek, McNamara, Birch, Bloom, & Heinrich, 2013; Cohen, 2007; Roazzi, Nyhof, & Johnson, 2013; Slingerland & Chudek, 2011).

Such explicit belief systems can, of course, be strictly philosophical positions on the mind–body problem. Mostly, however, they revolve around the proposed existence of a soul-like construct or another exclusively human property that survives bodily death (Bering, 2006; Boyer, 2001; see Anglin, 2014, for

lay-perceptions of souls vs. minds) or the existence of body-less minds that interact with the physical world. In fact, dualistic beliefs seem to be one of the prerequisites for the development of many elaborate supernatural beliefs, such as in a life after death (a mind without a body in the spiritual world), evil spirits or ghosts (a mind without a body in the physical world), or in reincarnation (a mind in a new body in the physical world) (e.g., Antony, 2006; Boyer, 2001; Bering, 2006; Bloom, 2007; Uhlmann et al., 2008). All of these beliefs require an individual to entertain the notion that mental states can somehow survive the death of a physical body, and therefore rely on endorsing the view that mental life is not fully explained by physical processes. Thus, considering that all humans seem to be natural-born dualists (Bloom, 2004), many scholars consider mind–body dualism one of the fundamental building blocks of more complex beliefs such as in gods and spirits, and ultimately for religious belief as a whole (Bloom, 2007). In other words, regardless of cultural background or educational influence, our tendency to view mind and body as distinct should produce similar belief systems in any given society. For one, a natural belief in minds that can exist without bodies (or non-human bodies), paired with our hard-wired hyperactive agency-detection and promiscuous teleology (Kelemen, 2004), should inadvertently lead to the formation of beliefs in gods or spirits (Bloom, 2007; Uhlmann et al., 2008). Additionally, in synergy with our inability to imagine the nonexistence of ourselves and others, motivational factors triggered by existential anxiety, as well as our constant search for meaning and purpose, a dualistic view on mind and body should also universally promote the formation of beliefs in some sort of afterlife (Bering, 2006; Uhlmann et al., 2008).

Supporting the theory that mind–body dualism and perspective-taking are deeply interlinked constructs and that mind–body dualism can be considered a prerequisite for religiosity, research found that mentalizing deficits—difficulties in acknowledging others' diverging mental content, as encountered, for example, in autism spectrum disorders (Frith & Happé, 1994)—are indeed associated with attenuated religious belief (Norenzayan, Gervais, & Trzesniewski, 2012). In one of the more complex studies on this topic thus far, Willard and Norenzayan (2013) investigated the relationship between the previously discussed psychological biases related to mind–body dualism in more detail, analyzing how strongly they predict supernatural and god beliefs. Consistently, self-reported explicit belief in mind–body dualism was the best predictor for both types of belief, with mind–body dualism, perspective-taking, teleology, and anthropomorphism revealing the expected positive intercorrelations. Similar results have been reported by Riecki, Lindeman, and Lipsanen (2013), who found that both strong dualistic and emergentistic beliefs predicted religious belief.

Recent work by Heflick, Goldenberg, Hart, and Kamp (2015) provided further empirical support for the proposition that mind–body dualism and afterlife beliefs are indeed related: afterlife beliefs were significantly heightened for participants under mortality salience—that is, after the inevitability of their death was made salient to them, typically causing existential anxiety (Greenberg et al., 1990)—but only when prompted to think of their selves as nonphysical in nature. Similarly,

work by Preston, Ritter, and Hepler (2013) found that rendering neurological explanations for the mind accessible decreased participants' belief in souls. Yet, if explanatory gaps in neuroscience were made salient, belief in souls increased, once more indicating that a default belief in dualism is indeed only suppressed by acquired scientific knowledge and is readily revived when opportunity arises.

As stated in the beginning of this section, the exact manifestations of explicit dualistic beliefs vary across cultures and times. While the concept of a soul in Abrahamic religions includes all human mental faculties (that is, one retains one's memories, emotions, and personality traits in the afterlife), Buddhist or Hindu reincarnation beliefs emphasize that one has no memories of one's former lives and only carries over a certain self-defining essence related to core personality traits (Smith, 1991). Other belief systems, such as animistic beliefs that are found in some indigenous tribal societies, include manifestations of dualistic beliefs that do not necessarily revolve around bodily death. Considered by Tyler (1871) to constitute the foundation of all religious belief, animism involves the attribution of a spiritual essence or a soul to nonhuman entities such as animals, rivers, trees, the wind, or fire. As another example for indigenous beliefs that are based on dualistic concepts, some Navajo believe in *yee naaldlooshii* [skin-walkers]—individuals who possess supernatural powers, including the ability to assume the physical form of another human or animal at will, without changing key elements of their mostly negative personality (Kluckhohn, 1944).

On the other hand, since the advent of secularization and advancements in neuroscience and related natural sciences, more and more people in Western societies reject the concept of an immortal soul (Lindeman, Riekkii, & Svedholm-Häkkinen, 2015), deny that mind and body are entirely independent entities (Ahn, Proctor, & Flanagan, 2009; Proctor, 2008), and have a slight tendency to agree with strictly physicalistic (i.e., materialistic monist) rather than dualistic statements (Hook & Farah, 2013). However, as research has shown that naïve implicit theories can in fact coexist with acquired scientific knowledge (Shtulman & Valcarcel, 2012), an explicit endorsement of physicalism must not necessarily mean that intuitions regarding mind–body relations match this direction. It's possible to imagine both a person who believes in a soul, yet has a strong feeling of *being* his or her body, and a person who explicitly believes that the mind is what the brain does, while feeling rather dissociated from his or her body. Although research on the interrelatedness between intuitive and explicit beliefs in mind—body dualism is scarce, some of our research found that both concepts might be positively, yet moderately related (Forstmann & Burgmer, 2015). Specifically, we found that the degree to which participants considered a physically duplicated hamster to have different mental states than its original counterpart positively correlated with an explicit measure of dualistic belief, in which participants used a pictorial item to indicate how they view the overlap between one's mind and body. This association is further confirmed by our aforementioned work on mind-body dualism and perspective-taking: people's tendency to engage in mental-state inference was affected by our manipulation of explicit beliefs in dualism, yet also correlated with a thought-experiment measure of intuitive dualistic beliefs

(Burgmer et al., 2016). In sum, previous work suggests that explicit and implicit beliefs in mind–body dualism are empirically related, yet conceptually distinct constructs of interest.

This theorizing can also help explain the phenomenon that philosopher Daniel Dennett refers to as the lay theory of a *Cartesian Theater* (Dennett, 1991). People in scientifically advanced societies oftentimes have a rather elaborate understanding of how the physical world interacts with the human organism. We all are taught in school how sound waves travel through the air into our ears, light sources emit photons that hit our retinas, and haptic sensations are caused by stimulation of nerve cells in our skin. Further, we all know that these inputs are translated into electric signals that travel along neural pathways all the way up into our brains where they are being processed. But what happens then? In the perception of many people, all these signals that have been processed, analyzed, and rearranged, ultimately arrive at some single point in the brain, where everything is collected and where then consciousness “happens”. That is, in this lay perception, although they acknowledge the fact that the brain is an important processor and analyzer of information, people believe that their selves reside at a certain single location in the brain where all information is subjectively “perceived” (i.e., the Cartesian theater), as if there were a small homunculus sitting there enthroned, who is presented what the eyes see and what the ears hear, and who is in charge of controlling everything.

However, scientists know now that this idea, although appealing, is a fundamental misconception, and that a Cartesian theater is in fact not really needed to explain our minds (e.g., Dennett & Kinsbourne, 1992). What we consider to be our self or consciousness is merely the sum of all the aforementioned processing and analyzing that the brain undertakes. The idea of a spatiotemporally unified self that can be located at a single point in the brain is in itself a modern form of mind–body dualism. In fact, Descartes himself assumed the soul to be connected to the pineal gland, which is—rather conveniently—located right in the center of the brain. This shows that even though people may explicitly subscribe to the idea that the mind is nothing more than neural activity, the previously discussed psychological biases and subjective experiences people make may lead them to intuitively continue to endorse a certain form of Cartesian mind–body dualism.

Consequences of Belief in Mind–Body Dualism

Finally, what implications does it have whether or not people explicitly endorse a dualistic view on minds and bodies? Similar to other beliefs and lay theories, such as beliefs in free will (e.g., Aarts & van den Bos, 2011; Alquist, Ainsworth, & Baumeister, 2013; Vohs & Schooler, 2008), beliefs in a just world (e.g., Callan, et al., 2009; Lerner, 1980) or beliefs in the malleability of personality traits (e.g., Dweck & Leggett, 1988; Molden & Dweck, 2006), beliefs in mind–body dualism can be assumed to have profound effects on people’s cognitions, emotions, and behaviors. While there is plenty of research on the effects of religious belief

(e.g., Shariff, Willard, Andersen, & Norenzayan, 2016), research into the isolated effect of viewing mind and body as separate is still surprisingly scarce. Some of our own recent experimental research revealed, for example, that dualistic beliefs are inversely related to health-related attitudes and behaviors (Forstmann et al., 2012). Participants whose dualistic beliefs were strengthened subsequently showed a greater disregard for their own personal health, both in the domains of attitudes (i.e., attitudes toward engaging in physical exercise, etc.) and actual behaviors (e.g., actual food consumption). Specifically, participants who read about mind–body dualism consumed more unhealthy food after participating in the experiment than did participants who read about physicalism. A possible explanation for these effects could be that a dualistic view on minds and bodies fosters the perception of one’s own body as a mere vessel or a tool, the primary function of which is to move the mind through space and to act as a sort of bridge between the mental and the physical world. As humans generally tend to value minds over bodies—in that they consider mentality to be the defining characteristic of a living human being (Gray, Gray, & Wegner, 2007; Gray, Knickman, & Wegner, 2011)—the view that body and mind are not one and the same may lead to people to engage more frequently in behaviors that momentarily provide mental pleasure while ultimately causing physical harm (e.g., consuming tasty, yet unhealthy food).

The negative relationship between dualistic beliefs and health behavior was further validated in a set of correlational field studies. People who had their lunch at a salad bar indicated a stronger belief in physicalism than did participants who had their lunch at a fast food burger joint. The same was true for participants who went shopping in a strictly organic (as opposed to “ordinary”) supermarket, and people who decided to take the stairs rather than the escalator at a metro station (Burgmer & Forstmann, 2016). Furthermore, in these studies, we were able to isolate a specific aspect of dualistic belief that seems to be responsible for the detrimental effects of dualism on health-related attitudes and behaviors. Particularly, participants who endorsed a dualistic view on mind–body relations were less likely to believe that bodily states influence mental well-being. This belief, in turn, was negatively related to people’s inclination to entertain health-sustaining attitudes or behaviors (Burgmer & Forstmann, 2016).

On the other hand, dissociating one’s mind from one’s body may also have potentially beneficial effects when dealing with traumatic experiences in which the body is subjected to harm. According to earlier literature, dissociation of mind and body can be “a defense often mobilized against the pain and helplessness engendered by traumatic experiences such as rape, incest, and combat. [...] It dissociates consciousness from the immediate experience of painful events: physical pain, fear, anxiety, helplessness” (Spiegel, 1986, pp. 123–124). Following a similar sentiment, newer clinical research suggests that dissociation from one’s body might indeed be a common defense mechanism in traumatic experiences: studies found that about 6% of the general population report a high number of dissociative experiences, with exceptionally higher levels among people with a history of childhood physical (and to a lesser extent sexual) abuse (Mulder, Beautrais, Joyce, & Fergusson, 1998), as well as a greater prevalence of dissociative symptoms in clinical patients with

childhood interpersonal trauma (Chu & Dill, 1990; Simeon et al., 2001). Further, some people report instances of dissociation “in which [their] self identity [became] detached from bodily sensation”(p. 460) while undergoing near-death experiences (Greyson, 2000), or show signs of depersonalization (a form of dissociation) in life-threatening situations (Noyes & Kletti, 1977). Thus, our intuitive mind–body dualism may be situationally strengthened to defend against harmful physical experiences, even to an extent that later manifests in chronic dissociative mental illness.

Another line of research by Thomas and Wardle (2014) found that aging can trigger a conflict between people’s still active minds and their weakening bodies, leading to a “defensive” increase in mind–body dualism with age. Highlighting the biopsychological effects the body can have on the mind, on the other hand, was found to promote a healthier lifestyle among older participants, who were eager to protect their minds as good as they can. Similarly related to the domain of health behavior, it was argued that a belief in mind–body dualism may increase stigma regarding mental health issues: framing the mind as independent from the body may attenuate the perception of mental disorders being a function of neurochemistry (Lebowitz, 2014; Miresco & Kirmayer, 2006). Thereby, responsibility for the condition may potentially be attributed to the patient, and a nonbiological view on mental disorders may promote prognostic pessimism. An empirical investigation by Kim, Ahn, Johnson, and Knobe (2016) found that lay theories about the origin of mental disorders indeed have profound real-life consequences: clinicians perceived mental disorders to be more biologically and less psychologically based when they were described in an abstract manner (i.e., by describing behavioral symptoms), and considered medication to be more effective in these cases, suggesting that a clinician’s view on minds and bodies may even affect the way he or she applies medical treatment.

Open Questions and Future Directions

Outside the domain of health behavior and religion, a belief in mind–body dualism may have effects on other common-sense beliefs or lay theories people hold about the world, either related to associated philosophical or psychological constructs. As stated earlier, a belief in mind–body dualism is presumably closely related to belief in free will. People seem to have a “transcendence vision” when it comes to how they explain the mind, in that they think people make decisions *based on* the mental states, yet do not consider mental states *causes* of these decisions (Knobe, 2014). In line with the idea of a Cartesian Theater, they view the self as something that “transcends all the states and processes within the mind, [...] the whole causal order” (p. 70), a dualistic view on mind and body that seems necessary for a belief in true free will. Not surprisingly, past research found substantial correlations between belief in mind–body dualism and free will (Nadelhoffer, Shepard, Nahmias, Sripada, & Ross, 2014). Causal evidence for this relation, however, is still lacking.

This causal evidence could help shed more light on the intimate relation between different metaphysical beliefs related to the philosophy of mind and further underline the important role that mind–body dualism plays as a potential root of some of these more sophisticated beliefs.

On a more psychological level, other beliefs and lay theories could be similarly affected by how much people dissociate their mind and body: one of these beliefs that pertain to people’s psychological workings is the belief in whether or not willpower is a limited resource. If people consider the mind to be independent of the body, they might consequently assume that bodily fatigue or other physiological states do not impact self-regulation capacities (a mental phenomenon) to the same extent as do people who feel a strong association between their mind and body. In other words, people who dissociate their mind from their body may believe to a lesser extent that the body affects the mind, and thereby adjust their belief in the limitedness of willpower accordingly. As past research points out, this shift in belief may even affect the *de facto* availability of self-regulation resources (Job, this volume; Job et al., 2010). Evidence for this link would provide additional support for the fact that even abstract metaphysical beliefs such as in mind–body dualism can alter fundamental and well-established psychological processes such as self-regulation.

With regard to the intuitive dualistic beliefs we discussed earlier, mind–body dualism may be causally related to other processes that are associated with the two processes we believe to be responsible for our dualistic intuitions, namely mental-state inference and bodily self-awareness. If dualists intuitively perceive a mind to be something that is separate from a body, they may take bodily displays of others less into consideration when they engage in perspective-taking, that is, when inferring their mental life. One could say, they might consider others’ bodily states to be less diagnostic of what is really going on in their mind. For example, a dualist may take another person’s body language or posture less into account when determining whether or not this person is telling a lie, whereas a physicalist (who believes in the unity of mind and body) may see the body as a “mirror” of the mind, and thus more strongly base his or her judgment on this observable information.

Conversely, the severed link between body and mind may lead dualists to take their own bodily states less into account when assessing how they themselves feel. This could as a consequence lower introspection accuracy, that is, it could negatively affect the degree to which the person is sensitive to aversive or pleasant bodily states. Thus, dualistic beliefs may in theory affect the very association we feel with our own bodies, thereby having a bearing on one of the most fundamental aspects of human self-perception.

Yet, a belief in the separation of one’s own mind and body may not just affect the perception of one’s body, but may also affect meta-perceptions of one’s mind. For example, the view that mental states are separate from the physical world and thereby inaccessible to anyone else may affect the degree to which people allow themselves to have socially inappropriate thoughts. If a physicalist believes mind and body to be one and the same, he or she may (nonconsciously) assume that mental states are in some way displayed by their body and are thus perceptible to

the outside world. Likewise, just as the use of deception is a marker for Theory of Mind development in children, the intuitive belief that our knowledge is uniquely accessible to ourselves—and not to others—may promote lying and deceptive behavior in adults. After all, if a person's true knowledge and feelings were separate from the observable world, it would make it quite a lot more difficult to catch a lie.

On the other hand, if this belief in the inaccessibility of minds also extends to other people, it may to a certain degree promote paranoid thinking patterns. For a dualist, a belief in the inaccessibility of other minds may foster the idea that others have goals and intentions that are different from what can be gathered from their overt behavior. Interestingly, paranoid thinking was already causally linked to mental-state inference (one of the presumed causes of dualistic thinking), in that it can be described as a hyperawareness of other's motives and intentions (Green & Phillips, 2004). Thereby, research on metaphysical beliefs such as in mind–body dualism may even extend to clinical research.

To conclude, future research will be needed to fully grasp the extent to which our intuitions and explicit belief regarding mind–body relations affect our beliefs, cognitions, emotions, and behavior. So far, little is known about which personality traits predict dualistic beliefs, how exactly situational variables affect lay theories on mind–body dualism, or whether there is more than two distinct kinds of belief in mind–body relations (e.g., emergentism, idealism). With the field of experimental philosophy of mind expected to continue growing over the next years, future research promises to substantially advance our understanding of these and other (philosophical) lay theories.

References

- Aarts, H., & van den Bos, K. (2011). On the foundations of beliefs in free will: Intentional binding and unconscious priming in self-agency. *Psychological Science*, *22*, 532–537.
- Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, *33*, 147–182.
- Alquist, J. L., Ainsworth, S. E., & Baumeister, R. F. (2013). Determined to conform: Disbelief in free will increases conformity. *Journal of Experimental Social Psychology*, *49*, 80–86.
- Anglin, S. M. (2014). I think, therefore I am? Examining conceptions of the self, soul, and mind. *Consciousness and Cognition*, *29*, 105–116.
- Antony, M. V. (2006). Simulation constraints, afterlife beliefs, and common-sense dualism. *Behavioral and Brain Sciences*, *29*, 462–463.
- Bal, M., & van den Bos, K. (2017). Effects of lay beliefs on the justice motive. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Bering, J. M. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, *29*, 453–462.
- Bering, J. M. (2011). *The belief instinct: The psychology of souls, destiny, and the meaning of life*. New York, NY: W. W. Norton & Company.
- Bering, J. M., & Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, *40*, 217–233.

- Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*, 382–386.
- Blackmore, S. (1987). Where am I? Perspectives in imagery and the out-of-body experience. *Journal of Mental Imagery*, *11*, 53–66.
- Blanke, O., & Arzy, S. (2005). The out-of-body experience: Disturbed self-processing at the temporo-parietal junction. *The Neuroscientist*, *11*, 16–24.
- Bloom, P. (2004). *Descartes' baby*. New York, NY: Basic Books.
- Bloom, P. (2007). Religion is natural. *Developmental Science*, *10*, 147–151. doi:10.1111/j.1467-7687.2007.00577.x
- Botvinick, M., & Cohen, J. (1998). Rubber hands “feel” touch that eyes see. *Nature*, *391*, 756. doi:10.1038/35784
- Boyer, P. (2001). *Religion explained*. New York, NY: Basic Books.
- Burgmer, P., & Forstmann, M. (2016). *Mind-body dualism and health revisited: Replications, mechanisms, and current findings from a field study*. Manuscript in preparation.
- Burgmer, P., Forstmann, M., Todd, A. R., & Mussweiler, T. (2016). Belief in mind-body dualism facilitates perspective-taking. *Cognition*. Manuscript under revision.
- Burnette, J. L., Hoyt, C. L., & Orvidas, K. (2017). Mindsets of body-weight. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Callan, M. J., Kay, A. C., Davidenko, N., & Ellard, J. H. (2009). The effects of justice motivation on memory for self- and other-relevant events. *Journal of Experimental Social Psychology*, *45*, 614–623.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*, 200–219.
- Chu, J. A., & Dill, D. L. (1990). Dissociative symptoms in relation to childhood physical and sexual abuse. *The American Journal of Psychiatry*, *147*, 887–892.
- Chudek, M., MacNamara, R., Birch, S. A. J., Bloom, P., & Henrich, J. (2013). *Developmental and cross-cultural evidence for intuitive dualism*. Manuscript in preparation.
- Cohen, E. (2007). *The mind possessed: The cognition of spirit possession in an Afro-Brazilian religious tradition*. Oxford, UK: Oxford University Press.
- Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy*, *68*, 87–106.
- Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little, Brown, & Co.
- Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, *15*, 183–201. doi:10.1017/S0140525X00068229
- Descartes, R. (1641). Meditations on first philosophy. In *The philosophical writings of René Descartes* (J. Cottingham, R. Stoothoff & D. Murdoch, Trans.). Cambridge, UK: Cambridge University Press, 1984.
- DeWall, C. N., MacDonald, G., Webster, G. D., Masten, C. L., Baumeister, R. F., Powell, C., et al. (2010). Acetaminophen reduces social pain: Behavioral and neural evidence. *Psychological Science*, *21*, 931–937.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.
- Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science*, *317*, 1048.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*, 864–886. doi:10.1037/0033-295X.114.4.864
- Estes, D. (1994). Young children’s understanding of the mind: Imagery, introspection, and some implications. *Journal of Applied Developmental Psychology*, *15*, 529–548.
- Forstmann, M., & Burgmer, P. (2015). Adults are intuitive mind-body dualists. *Journal of Experimental Psychology: General*, *144*, 222–235.
- Forstmann, M., Burgmer, P., & Mussweiler, T. (2012). The mind is willing, but the flesh is weak: The effects of mind-body dualism on health behavior. *Psychological Science*, *23*, 1239–1245.
- Frith, U., & Happé, F. (1994). Autism: Beyond “theory of mind”. *Cognition*, *50*, 115–132.

- Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8, 404–409.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*, 7, 287–292.
- Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56, 165–193.
- Gervais, W. M., & Norenzayan, A. (2012). Analytic thinking promotes religious disbelief. *Science*, 336, 493–496.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Boston, MA: MIT Press.
- Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, 7, 145–171.
- Gottfried, G. M., Gelman, S. A., & Schultz, J. (1999). Children's understanding of the brain: From early essentialism to biological theory. *Cognitive Development*, 14, 147–174.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- Gray, K., Knickman, T. A., & Wegner, D. M. (2011). More dead than dead: Perceptions of persons in the persistent vegetative state. *Cognition*, 121, 275–280.
- Green, M. J., & Phillips, M. L. (2004). Social threat perception and the evolution of paranoia. *Neuroscience and Biobehavioral Reviews*, 28, 333–342.
- Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., et al. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology*, 58, 308–318.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4–27.
- Greyson, B. (2000). Dissociation in people who have near-death experiences: Out of their bodies or out of their minds? *Lancet*, 355, 460–463.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2, 271–299.
- Guterstam, A., Petkova, V. I., & Ehrsson, H. H. (2011). The illusion of owning a third arm. *PLoS ONE*, 6, e17208.
- Guthrie, S. E. (1993). *Faces in the clouds: A new theory of religion*. New York, NY: Oxford University Press.
- Hala, S., Chandler, M., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of three-year-olds' understanding of false belief. *Child Development*, 62, 83–97.
- Hamlin, J. K., Hallinan, E. V., & Woodward, A. L. (2008). Do as I do: 7-month-old infants selectively reproduce others' goals. *Developmental Science*, 11, 487–494.
- Heflick, N. A., Goldenberg, J. L., Hart, J., & Kamp, S. M. (2015). Death awareness and body-self dualism: A why and how of afterlife belief. *European Journal of Social Psychology*, 45, 267–275.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57, 243–259.
- Hood, B. M., & Bloom, P. (2008). Children prefer certain individuals over perfect duplicates. *Cognition*, 106, 455–462.
- Hood, B., Gjersoe, N., & Bloom, P. (2012). Do children think that duplicating the body also duplicates the mind? *Cognition*, 125, 466–474.
- Hook, C. J., & Farah, M. J. (2013). Look again: Effects of brain images and mind-brain dualism on lay evaluations of research. *Journal of Cognitive Neuroscience*, 25, 1397–1405.
- Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., et al. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron*, 70, 363–374.
- Job, V. (2017). Effects of lay theories on self-control. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories-how beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion-Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science*, 21, 1686–1693.

- Johnson, C. N. (1990). If You Had My Brain, Where Would I Be? Children's understanding of the brain and identity. *Child Development, 61*, 962–972.
- Johnson, C. N., & Wellman, H. M. (1982). Children's developing conceptions of the mind and brain. *Child Development, 53*, 222–234.
- Kafka, F. (1915). *The metamorphosis and other stories* (p. 1996). New York, NY: Barnes & Noble.
- Kelemen, D. (2004). Are children "intuitive theists"? Reasoning about purpose and design in nature. *Psychological Science, 15*, 295–301.
- Kelley, H. H. (1967). Attribution theory in social psychology. In *Nebraska symposium on motivation*. Lincoln, NE: University of Nebraska Press.
- Kim, J. (2000). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge, MA: The MIT Press.
- Kim, N. S., Ahn, W. K., Johnson, S. G., & Knobe, J. (2016). The influence of framing on clinicians' judgments of the biological basis of behaviors. *Journal of Experimental Psychology: Applied, 22*, 39–47.
- Gluckhohn, C. (1944). *Navaho witchcraft*. Boston, MA: Beacon Press.
- Knobe, J. (2007). Experimental philosophy. *Philosophy Compass, 2*, 81–92.
- Knobe, J. (2014). Free will and the scientific vision. In E. Machery & E. O'Neill (Eds.), *Current controversies in experimental philosophy*. London, UK: Routledge.
- Kuhlmeier, V. A., Bloom, P., & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition, 94*, 95–103.
- Lebowitz, M. S. (2014). Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice, 21*, 67–83.
- Legerstee, M. (1992). A review of the animate-inanimate distinction in infancy: Implications for models of social and cognitive knowing. *Early Development and Parenting, 1*, 59–67.
- Lengenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: Manipulating bodily self-consciousness. *Science, 317*, 1096–1099.
- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Lillard, A. S. (1996). Body or mind: Children's categorizing of pretense. *Child Development, 67*, 1717–1734.
- Lindeman, M., Riekkki, T., & Svedholm-Häkkinen, A. M. (2015). Individual differences in conceptions of soul, mind, and brain. *Journal of Individual Differences, 36*, 157–162.
- Marshall, P. J., & Comalli, C. E. (2012). Young children's changing conceptualizations of brain function: Implications for teaching neuroscience in early elementary settings. *Early Education & Development, 23*, 4–23.
- Miresco, M. J., & Kirmayer, L. J. (2006). The persistence of mind-brain dualism in psychiatric reasoning about clinical scenarios. *American Journal of Psychiatry, 163*, 913–918.
- Molden, D. C., & Dweck, C. S. (2006). Finding "meaning" in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist, 61*, 192–203.
- Mulder, R. T., Beautrais, A. L., Joyce, P. R., & Fergusson, D. M. (1998). Relationship between dissociation, childhood sexual abuse, childhood physical abuse, and mental illness in a general population sample. *The American Journal of Psychiatry, 155*, 806–811.
- Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., & Ross, L. T. (2014). The free will inventory: Measuring beliefs about agency and responsibility. *Consciousness and Cognition, 25*, 27–41.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*, 435–450.
- Nahmias, E., Shepard, J., & Reuter, S. (2014). It's OK if 'my brain made me do it': People's intuitions about free will and neuroscientific prediction. *Cognition, 133*, 502–516.
- Nilsson, J. P., Söderström, M., Karlsson, A. U., Lekander, M., Åkerstedt, T., Lindroth, N. E., et al. (2005). Less effective executive functioning after one night's sleep deprivation. *Journal of Sleep Research, 14*, 1–6.
- Norenzayan, A., Gervais, W. M., & Trzesniewski, K. H. (2012). Mentalizing deficits constrain belief in a personal god. *PLoS ONE, 7*, 1–8.
- Noyes, R., & Kletti, R. (1977). Depersonalization in response to life-threatening danger. *Comprehensive Psychiatry, 18*, 375–384.
- Parfit, D. (1984). *Reasons and persons*. Oxford, UK: Oxford University Press.

- Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science, 11*, 353–359.
- Pilcher, J. J., & Huffcutt, A. J. (1996). Effects of sleep deprivation on performance: A meta-analysis. *Sleep: Journal of Sleep Research & Sleep Medicine, 19*, 318–326.
- Plato. (360 B.C.). *Timaeus and Critias*, (Desmond Lee, Trans., Ed.). London, UK: Penguin, 1977.
- Povinelli, D. J., & Bering, J. M. (2002). The mentality of apes revisited. *Current Directions in Psychological Science, 11*, 115–119.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515–526.
- Preston, J. L., Ritter, R. S., & Hepler, J. (2013). Neuroscience and the soul: Competing explanations for the human experience. *Cognition, 127*, 31–37.
- Preston, J., Gray, K., & Wegner, D. M. (2006). The Godfather of soul. *Behavioral and Brain Sciences, 29*, 482–483.
- Proctor, C. C. T. (2008). *Clinicians' and laypeople's beliefs about the causal basis and treatment of mental disorders*. New Haven: Yale University. (unpublished dissertation).
- Rattan, A., & Georgeac, O. (2017). Mindsets about malleability and intergroup relations. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories-how beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Riekki, T., Lindeman, M., & Lipsanen, J. (2013). Conceptions about the mind-body problem and their relations to afterlife beliefs, paranormal beliefs, religiosity, and ontological confusions. *Advances in Cognitive Psychology, 9*, 112–120.
- Roazzi, M., Nyhof, M., & Johnson, C. (2013). Mind, soul and spirit: Conceptions of immaterial identity in different cultures. *International Journal for the Psychology of Religion, 23*, 75–86.
- 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.
- Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., & Lopez, C. (2013). Bodily ownership and self-location: Components of bodily self-consciousness. *Consciousness and Cognition, 22*(4), 1239–1252.
- Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2016). Religious priming: A meta-analysis with a focus on prosociality. *Personality and Social Psychology Review, 20*, 27–48. doi:10.1177/1088868314568811
- Shulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124*, 209–215.
- Simeon, D., Guralnik, O., Schmeidler, J., Sirof, B., & Knutelska, M. (2001). The role of childhood interpersonal trauma in depersonalization disorder. *American Journal of Psychiatry, 158*, 1027–1033.
- Slingerland, E., & Chudek, M. (2011). The prevalence of mind-body dualism in early China. *Cognitive Science, 35*, 997–1007.
- Slooman, S. (2009). *Causal models: How people think about the world and its alternatives*. Oxford, UK: Oxford University Press.
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*, 946–958.
- Smith, H. (1991). *The world's religions*. New York: HarperCollins Publishers.
- Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essay on biology and cognition* (pp. 133–169). Hillsdale, NJ: Lawrence Erlbaum.
- Spiegel, D. (1986). Dissociating damage. *The American Journal of Clinical Hypnosis, 29*, 123–131.
- Starmans, C., & Bloom, P. (2012). Windows to the soul: Children and adults see the eyes as the location of the self. *Cognition, 123*, 313–318.
- Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition, 124*, 272–283.
- Stucke, T. S., & Baumeister, R. F. (2006). Ego depletion and aggressive behavior: Is the inhibition of aggression a limited resource? *European Journal of Social Psychology, 36*, 1–13.

- Suddendorf, T., & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes, and other animals. *Psychological Bulletin*, *127*, 629–650.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.
- Thomas, M., & Wardle, G. A. (2014). “I’m more aware of the body than ever”: Old Adults’ experiences of the psychosocial significance of their bodies. *Psychological Studies*, *59*, 11–21.
- Tyler, E. B. (1871). *Primitive culture: Researches into the development of mythology, philosophy, religion, language, art, and custom*. London, England: John Murray.
- Uhlmann, E. L., Poehlman, T. A., & Bargh, J. A. (2008). Implicit theism. In R. M. Sorrentino & S. Yamaguchi, *Handbook of motivation and cognition across cultures* (pp. 71–94). Oxford, UK: Elsevier.
- Vohs, K. D., & Schooler, J. W. (2008). The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science*, *19*, 49–54.
- Watson, J. K., Gelman, S. A., & Wellman, H. M. (1998). Young children’s understanding of the non-physical nature of thoughts and the physical nature of the brain. *British Journal of Developmental Psychology*, *16*, 321–335.
- Wegener, D. T., & Petty, R. E. (1998). The naive scientist revisited: Naive theories and social judgment. *Social Cognition*, *16*, 1–7.
- Wellman, H. M., & Estes, D. (1986). Early understanding of mental entities: A reexamination of childhood realism. *Child Development*, *57*, 910–923.
- 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.
- Willard, A. K., & Norenzayan, A. (2013). Cognitive biases explain religious belief, paranormal belief, and belief in life’s purpose. *Cognition*, *129*, 379–391.
- 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.
- Woodward, A. (1998). Infants selectively encode the goal object of an actor’s reach. *Cognition*, *69*, 1–34.
- Woodward, A. L. (1999). Infants’ ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, *22*, 145–160.
- Woolley, J. D., & Wellman, H. M. (1992). Children’s conceptions of dreams. *Cognitive Development*, *7*, 365–380.

Lay Theories of the Mind/Brain Relationship and the Allure of Neuroscience

Diego Fernandez-Duque

Allure: the quality of being powerfully and mysteriously attractive or fascinating.

The '90s were declared the Decade of the Brain by United States' President George H. W. Bush, who at the time of his proclamation reflected that "the human brain, a 3-pound mass of interwoven nerve cells that controls our activity, is one of the most magnificent and mysterious wonders of creation. The seat of human intelligence, interpreter of senses, and controller of movement, this incredible organ continues to intrigue scientists and layman alike" (Library of Congress' Website, <https://www.loc.gov/loc/brain/>). Upon reading those sentences, it is hard not to feel some amount of sympathy for the author, as he struggles to find the words to best describe the relation between brain and mind. Absent from the proclamation are the attributes we most often think of as uniquely human: a notion of self, rational thought, language, free will. Similarly, no reference is made about cognition, consciousness, emotion, or the mind. Instead, the president plays it safe, mentions *intelligence*, focuses on the brain's ability for sensation decoding and movement control, and calls it a day.

In all fairness, most people probably could not have done much better. Philosophers have been debating the mind/brain question for centuries. In the 1600s, Descartes proposed that humans were a combination of body and mind (Descartes, 1984/1641). For Descartes, the body was part of the natural world and as such it was bound by the laws of nature. But the mind, Descartes thought, was capable of abilities that were uniquely human, including moral evaluation, appreciation of beauty, and free will. Philosophers have come a long way since the days of Descartes, and his *substance dualism* has run out of favor in most philosophical circles. However, it is unclear where ordinary folk stand on the issue. Many religious beliefs that are popular across the globe, such as beliefs in the afterlife and in

D. Fernandez-Duque (✉)

Psychology Department, Villanova University,
800 Lancaster Ave, Villanova, PA 19085, USA
e-mail: diego.fernandezduque@villanova.edu

the existence of the soul, necessitate a dualist concept of mind/brain relation (Bering & Bjorklund, 2004). Most people believe both in free will and moral responsibility (Monroe & Malle, 2010; Nahmias, Morris, Nadelhoffer, & Turner, 2005), holding others responsible for actions only in situations where a choice to act differently was available; the dualist view can easily accommodate such a perspective. On the other hand, the current explosion of research in neuroscience, with descriptions in the popular media of brains doing things that minds were supposed to do (Racine, Bar-Ilan, & Illes, 2005), has perhaps begun to challenge the beliefs that ordinary people have inherited from Descartes. For researchers interested in studying people's *common sense* theories about brain and mind, the challenge ahead is to find methods that faithfully capture those lay theories.

How to Study Lay Theory of Mind/Brain Relation?

Several approaches exist to probe lay theories of mind/brain, each of which has its own sets of strengths and weaknesses. One option is to ask people directly about their beliefs. This is sometimes done in the form of questionnaires probing different variants of dualism (e.g., mind and body are qualitatively distinct), materialism (mind and body are the same or fundamentally united), emergentism (mind and brains are qualitatively different but interdependent), or some other form of—*isms* (Demertzi et al., 2009; Stanovich, 1989). Questionnaires allow researchers to explore possible associations with religious beliefs, and with other folk beliefs like belief in the afterlife and belief in free will. Other times, researchers ask simple questions—such as “do you need a brain to think?”—in order to assess people's common sense beliefs. An obvious strength of this approach is its face validity. If we want to know what people think about X, asking them directly what they think about X seems a sensible first step. But this approach risks running into problems if people are inconsistent, both in the sense that their answers might differ depending on how they are probed, and in the sense that their reflective answer might not align with the set of background beliefs they regularly hold for judgments in everyday life. Answers to questionnaires may also be susceptible to educational biases and cultural influences. Depending on the research question, such variability may be seen as a strength or a weakness.

To overcome these limitations, lay theory researchers sometimes probe participants' knowledge in more indirect ways. Instead of asking people to reflect on the nature of the mind/brain relation—a prospect overloaded with philosophical baggage—people may be asked to entertain more mundane scenarios designed to assess their beliefs. For example, they may be asked to consider a certain neurological disease, as in the case of frontotemporal dementia, and predict its psychological consequences (Strohmingner & Nichols, 2015). They may be asked to assess how much the brain contributes to various mental constructs, including personality traits,

cognitive processes, and even the sense of self (Fernandez-Duque & Schwartz, 2016; Johnson & Wellman, 1982). By breaking the mind/brain problem into smaller units, researchers aim to obtain a more exact assessment of people's lay beliefs.

Finally, another approach is to probe mind/brain beliefs implicitly. In other words, rather than asking people about the relation between mind and brain, one could ask people to perform an ostensibly unrelated judgment that offers a window into their lay beliefs. For example, one could ask participants to assess the quality of an explanation of a psychological phenomenon, as a way to test whether the presence of neuroscience information increases the perceived quality of the explanation (it does; Fernandez-Duque, Evans, Christian, & Hodges, 2015; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008). Alternatively, one could ask participants to assess the moral responsibility behind hypothetical moral transgressions, as a way to test whether providing neuroscience information increases deterministic explanations of behavior and reduces moral condemnation (it does; Monterosso, Royzman, & Schwartz, 2005). In sum, this implicit approach uses neuroscience information—that is, information about the brain—as the independent variable to explore its influence on judgments of mental life. To the extent that those influences are documented in research, they provide psychologists with evidence (albeit of the indirect variety) that a connection exists between brain and mind in people's lay theories. Obviously, there is a big explanatory gap between showing that a connection exists and providing a full description of such a relation. This limitation notwithstanding, the study of the neuroscience allure has sparked interest in its own right, due to its relevance for many disciplines, including mental health, education, and law.

Overview of the Chapter

It is the last approach—the implicit assessment of beliefs about mind and brain—that constitutes the bulk of this chapter. Particular emphasis is given to the allure of neuroscience explanations. Next, I review the effect of neuroscience on judgments of responsibility and free will, as studied in the fields of psychology and law. This is followed by a review of folk beliefs on brain/mind/self, and a brief review of the developmental literature on this issue. I conclude the chapter by pointing out the importance of research on lay theories of mind/brain, arguing that different conceptualizations of the mind/brain relation have profound implications for public policy in mental health research and practice.

The Allure of Neuroscience: A Brief History

Over the last quarter century, our understanding of the mind and the brain have undergone a revolution. At the center of those changes has been the development of new neuroimaging techniques that have allowed scientists to create maps

connecting mental processes to their putative neural substrates. The mapping process is inferential and overwrought with statistical assumptions, but maps they are, pictures of brain activity where years ago there was only mind. Importantly, many of the new techniques are noninvasive, and therefore safe for use in humans. As a consequence, those uniquely human capacities that so long fascinated Descartes have, in the last two decades, become ripe for neuroscientific inquiry: the neural mechanisms of creativity, rational thought, morality, language, and the self are no longer out of bounds for scientific exploration.

The scientific advances of the last three decades have been accompanied by increased attention from the media and increased fascination by the public. The attention is deserved; cognitive neuroscience has provided great additional explanatory power to the mechanisms that underlie psychological processes. Sometimes, however, superfluous information is added that does not provide additional insight. What happens then? Are people fooled by cognitive neuroscience, or by the images we usually associate with it?

To start answering this question, McCabe and Castel (2008) had participants read a one-page summary of a cognitive neuroscience finding written for the popular press. This baseline condition was compared to experimental conditions in which the same neuroscience information was accompanied by either a functional MRI image or a bar chart. Participants rated the scientific reasoning most highly when the neuroscience explanation was paired with the fMRI. McCabe and Castel concluded that brain images conferred credibility to the neuroscience explanations. However, subsequent studies have failed to replicate these findings, and the current consensus in the field is that brain images have little to no effect on the perceived quality of neuroscientific explanations. However, it remains a possibility that neuroscience information—pictorial or text-based—might influence the perceived quality of psychological explanations. We turn to that literature next.

Neuroscience Increases the Appeal of Psychological Explanations

What happens when dubious references to brain mechanisms are brought up to pseudo-explain a psychological phenomenon? In those instances, are audiences more accepting of neuro-gibberish than of regular gibberish, and if so, why? To start answering this question, Weisberg and collaborators asked participants to read vignettes about well-established psychological phenomena and their possible explanations (Weisberg et al., 2008). The description of the phenomena was always accurate, but the quality of the explanation was variable: sometimes the arguments were good while other times they were circular, a mere restatement of the phenomenon. Consider, for example, a vignette reporting that, in visual tasks, the patterns of response times were different for faces than for places. In that vignette, the good explanation said this happened “because people use different processes to

recognize faces than they use to recognize places.” In contrast, the circular explanation claimed it happened “because the participants’ responses were contingent on whether they saw a face or a place on the screen”. A second factor provided the critical manipulation: half of the vignettes included superfluous neuroscience sentences, while the other half did not. The neuroscience information was not wrong, it was simply irrelevant to the phenomenon it was trying to explain. In the above example, it said that “neuroscientists have shown that the extrastriate cortex—an area of the brain known to be involved in processing complex visual stimuli—is activated by pictures of faces and places.” Despite its lack of relevance, such superfluous neuroscience sentences increased the perceived quality of circular explanations.

Weisberg’s findings have been largely replicated by other labs, confirming that the ‘allure of neuroscience’ is conceptual rather than perceptual, meaning that neuroscience information is persuasive regardless of whether it is presented in form of brain images or neuroscientific text (Fernandez-Duque et al., 2015; Michael, Newman, Vuorre, Cumming, & Garry, 2013). This raises the question: why is neuroscience so alluring? One answer might be that neuroscience is a prestigious science that people trust. A different answer would be that neuroscience offers reductive explanations of psychological phenomena.

Unlike the social sciences, neuroscience is considered a ‘real’ science. When asked about the prestige of neuroscience, or about the gap between a neuroscience expert and a lay person, or about the scientific rigor of the discipline, undergraduate students always cluster neuroscience with other natural sciences and away from social sciences and psychology (Fernandez-Duque et al., 2015). Other studies show that people believe that biological explanations are more complex and more scientific than psychological explanations. This bias toward the natural sciences emerges as early as kindergarten, and vestiges of it can be observed in adulthood (Keil, Lockhart, & Schlegel, 2010). For example, the mere presence of a nonsense math equation increases the perceived quality of a scientific abstract (Eriksson, 2012), and the inclusion of a chemical formula increases the belief in a medication’s efficacy (Tal & Wansink, 2014); just telling people that scientists understand a phenomenon is enough to increase people’s judgment of their own understanding (Sloman & Rabb, 2016). In sum, there is little doubt that neuroscience is held in high regard as a science, and that scientific jargon often creates an illusion of understanding. Put these two facts together, and one might conclude that the allure of neuroscience is driven by its prestige.

However, if neuroscience’s allure in explanations of psychological phenomena had to do with its status as a prestigious science, we would expect that gibberish from other “hard sciences” would also be alluring, provided that its relation to the psychological phenomena was not too far fetched; after all, such hard sciences—unlike psychology—are prestigious too. In our study, we explored this hypothesis by including irrelevant information from the hard sciences; for example, in the previously described vignette on visual processing of faces and places, participants read that “computational scientists have used spectrograms to show that pictures of faces and places convey a range of spatial frequencies.” As in the case of

neuroscience, the information was true but not particularly informative in explaining the psychological finding (Fernandez-Duque et al., 2015). The results of the study show convincingly that pseudo-explanations from other hard sciences are not as compelling as neuroscience pseudo-explanations (Fernandez-Duque et al., 2015).

What, then, explains the neuroscience allure? Although a definitive answer to this question has not yet been reached, we and others have speculated that the reason why neuroscience information is seen as more relevant than hard science or psychological counterparts is that in Western cultures, educated people conceptualize the brain as the engine of the mind (Fernandez-Duque & Schwartz, 2016; Hopkins, Weisberg, & Taylor, 2016). That is, participants in our studies may be conceptualizing the brain as the physical substrate that instantiates the mind, the structure to which psychological phenomena may one day be mapped or reduced.

The perspective we have been advocating so far in this section is that the allure of neuroscience for psychological phenomena stems from a lay theory according to which the brain is the engine of the mind, or put slightly differently, it is the next level of analysis below the mind. If this perspective is correct, then the allure of neuroscience would be just a special case of a more general heuristic, by which information is alluring if it sits at the level below the phenomenon of interest. According to this position, the most alluring information is reductive information. To test this hypothesis, Hopkins et al. (2016) extended the original paradigm to include other scientific disciplines, such as physics, chemistry, and biology. As predicted, they found that superfluous information was most alluring at the level just below the phenomena in need of explanation. In other words, people did prefer reductive explanations. As a caveat, it should be pointed out that the allure for the psychology/neuroscience pair was larger than for any other pair, leaving open the possibility of additional content specific influences above and beyond the allure of reductive explanations.

Neuroscience Influences Judgments of Responsibility

In 1848, a railroad worker named Phineas Gage suffered a terrible accident when a metal rod exploded in his face, impaled him through cheek bone and skull, and in its way destroyed large parts of his frontal lobe. Gage survived, but his personality changed profoundly. Previously, he had been a conscientious worker, intelligent and well adapted. After the accident, he became “fitful, irreverent, indulging at times in the grossest profanities (which was not previously his custom), manifesting little deference for his fellows, impatient of restraint or advice when it conflicts with his desires” (Harlow, 1868). Friends and acquaintances would report that ‘Gage was no longer Gage’. Most of us would not blame Gage. Instead, we would assume that his behavioral outbursts and moral transgressions were outside of his control—that he did not have much choice—and conclude that he should not be held responsible, or at least not as responsible as someone with full mental capacity.

Probably, we would not need to see a picture of his brain or his skull to reach these conclusions.

Our response to Gage's behavior nicely illustrates some of the folk beliefs regarding brain, moral responsibility, and free will. Over the last decade, researchers have begun to systematically assess such folk beliefs (Greene & Cohen, 2004). These studies are remarkably consistent in showing that neuroscience information does influence participants' judgments of moral responsibility. In other words, this literature on neuroscience and responsibility tells the very same story already presented regarding the allure of neuroscience and psychological phenomena.

In one of the first studies of its kind, Monterosso, Royzman, and Schwartz (2005) asked participants to read vignettes describing individuals who had committed a moral transgression. The authors varied whether the explanation for the transgression was neurobiological (e.g. "unusually high levels of a particular neurotransmitter") or experiential (e.g., "severely and brutally abused as a child"), reasoning that a neurobiological explanation would elicit a mechanistic view, whereas an experiential one would not. As predicted, the neurobiological explanation led to less blame than the experiential explanation, and to reduced ascription of willful control. In a follow-up study, the same two explanation types (neurobiological, experiential) were factorially crossed with the presence or absence of a neuroimage. Participants were asked to judge the extent to which the transgression was due to lack of moral character. By itself, the experiential justification led to larger moral condemnation than the neurobiological justification, but this difference disappeared when a neuroimage was attached. Participants who saw the brain image together with the experiential explanation probably took the brain to be the mechanistic mediator of the experiential account, and thus reduced the target's responsibility (Beall et al., 2013).

Similar findings have been obtained in experiments in the field of psychology and law. For example, when participants in a mock trial had to decide on a case of not guilty by reason of insanity, they had a tendency to find neuroscience-based evidence more persuasive than psychological evidence (Schweitzer & Saks, 2011). In another study, the presence of neuroscientific testimony reduced the likelihood of a death sentence verdict; the reduction occurred irrespective of whether the evidence consisted of a brain image or neuropsychological testimony alone (Greene & Cahill, 2012). In yet another study, participants in a mock trial had to decide on a case of not guilty by reason of insanity. Participants were biased toward a not guilty verdict by evidence of neuroscience lesion (in this case, an MRI image), as well as by evidence of abrupt onset (in this case, an episode of traumatic brain injury) (Gurley & Marcus, 2008). Both of these factors—neural evidence and abrupt onset—are also evident in the illustrative case of Phineas Gage. Finally, one study using US state judges as participants showed that expert testimony on the neurobiological mechanisms of psychopathy causes judges to consider those mechanisms as mitigating factors, thus leading to reduced criminal sentences (Aspinwall, Brown, & Tabery, 2012).

In summary, reading or listening to evidence about the neural bases of human behavior leads people away from attributions of moral responsibility and away from retributive punishment. One account of these results is that learning about neuroscience highlights a mechanistic worldview in which free will is diminished and therefore actors should be held less blameworthy for their acts (Greene & Cohen, 2004; Shariff et al., 2014). That account rests in part on the assumption that ordinary folk deem neuroscientific explanations of behavior to be constraints on free will. Whether this is indeed the case is a matter of debate, which is discussed next.

Neuroscience Influences Judgments of Free Will

In the philosophical literature, ‘free will’ is discussed by appeal to metaphysical concepts such as ‘uncaused agency’ (i.e., the ability of an agent to act without such act being caused by something else). Free will thus defined is challenged by a deterministic world, and this leads philosophers to all sorts of intellectual contortions to try to establish a coherent view (Roskies, 2006).¹ Research exploring the relation between neuroscience and free will philosophically defined has had limited success; there is great variability in people’s judgments across experiments, and there are often internally inconsistent responses. For example, some studies suggest that people embrace both determinism *and* free will, a position known as *compatibilism* (Monroe & Malle, 2010; Nahmias, 2006; Nahmias, Shepard, & Reuter, 2014; Nichols & Knobe, 2007). According to this view, when morally evaluating an action, people state that even if the universe is fully deterministic, the actor could act differently. Other studies show instead that when presented with neuroscience claims that “free will does not exist because choices are caused by neural impulses” people reply by appealing to a different level of analysis, focusing on the agent to argue that “the person makes the neural impulses happen.” This way, people seem to endorse the neuroscientific correlates of psychological states without committing to a deterministic view of them. Yet some other research suggests that determinism undermines free will in the abstract, but does not excuse wrongdoing in concrete cases (Nichols, 2011). In general, the sense one gets from reading this literature is that people lack stable notions of ‘free will’.

As it turns out, ‘free will’ as understood by the common person is quite a different concept from the one developed by professional philosophers. When asked to define ‘free will’, ordinary folk do not refer to metaphysical criteria, rather, they provide a psychological account. People report that ‘free will’ consists of the ability to make choices consistent with one’s desires, reasonably free of constraints; they sometimes also emphasize the reflective, deliberate aspect of it, that is, the

¹These contortions include rejecting determinism to save free will (libertarians); conceding that the world is deterministic and thus acknowledging that that free will does not exist (hard determinists); or accepting determinism but still claim that free will is possible (i.e., compatibilism).

forethought of weighting the pros and cons of the action (Monroe & Malle, 2010; Nahmias, 2016).

How does the folk concept of ‘free will’ relate to neuroscience and morality? The answer, at least hypothetically, is remarkably simple and powerful. To the extent that a neurological disorder disrupts one of the underlying psychological components of ‘free will’ (choice, desire, absence of external constraint, forethought), a reduction of ‘free will’ will ensue. In contrast, the existence of a normally functioning brain ought not pose a challenge to free will because a normal brain—by definition—has a correspondingly normal psychology. The threat of determinism that so much challenges the concept of ‘free will’ metaphysically defined simply vanishes once we adopt its common sense definition (Nahmias, 2006).

So far, I have described two competing conceptualizations of free will: one favored by philosophers which is metaphysically defined, and the other one favored by folk theory which is psychologically defined by ‘choice’. What role do these two different conceptualizations of ‘free will’ play in judgments of moral responsibility? One way to answer this question is to pit the metaphysical notion of free will against the folk notion, in a 2×2 factorial design. In one such study, participants were divided into an experimental group that read a statement arguing against metaphysical free will (“all behavior is determined by brain activity, which in turn is determined by a combination of environmental and genetic factors”) and a control group that read a neutral statement that made no reference to free will (“Oceans cover 71% of the earth’s surface”). As expected, the experimental group reported less belief in metaphysical free will than the control group. Immediately after this manipulation, the participants watched a brief video in order to make a judgment of blame. In the ‘choice’ condition, the video depicted a person in a situation in which he could choose to steal money from a partner. In contrast, in the ‘no choice’ condition the amount of money taken was determined randomly. Participants who saw the ‘choice’ video assigned much more blame than participants in the ‘no choice’ condition. Importantly, the metaphysical manipulation had no effect (Monroe, Brady, & Malle, 2016). In another study, a vignette described a hypothetical study in which scientists were able to predict a person’s future behavior based on her pattern of neural activation. Despite the scientists being able to predict in advance what the person would do (consistent with determinism), participants thought that the person was still exercising her free will (Nahmias et al., 2014). Only if the vignette described neuroscientists as bypassing the person’s decision—that is, if it described the scientists as stimulating the patient’s brain to actively manipulate her choice—did participants consider the patient to be deprived of free will.

In sum, the folk judgment of whether a person has ‘free will’ seems to depend on the mental states of the person: if her desires and choices are efficacious for causing action, then free will is affirmed, and moral responsibility assigned. This framework is useful in helping us reinterpret the literature on neuroscience and judgments of responsibility. We excuse Phineas Gage’s moral transgressions not because our knowing of his brain lesion turns us into hard determinists skeptical of metaphysical

free will, but rather because our knowing of his lesion reminds us of Gage's reduced mental capacities to make choices consistent with his desires and free of unreasonable constraints.

Alternatively, and on a more speculative note, when it comes to commitments regarding the relationships among mind, brain, and free will, people may be, in the words of philosopher Eddie Nahmias, 'theory-lite' (Nahmias, 2017); that is, people may have intuitions that are unstable and/or contradictory, without a reliable, strong commitment to dualism or materialism. After all, unlike professional philosophers, people can get through life without having to address metaphysical questions of the mind/brain relation. On the other hand, people cannot get through life without some form of moral theory to guide their judgments of blame and responsibility. Those judgments are dependent on free will in the psychological sense of the term. As neuroscience progresses, it seems likely that people will retain their belief-based model of free will and morality, and simply make ad hoc necessary adjustments to their 'theory-lite' metaphysics for the rare occasions in which such esoteric questions may arise.

Interlude

Up to this point, the chapter has focused on the implicit assessment of beliefs about mind/brain relation. I have described the allure of neuroscience for psychological explanations, and the effect of neuroscience on judgments of responsibility and free will. The approach has been successful in showing that appeals to neuroscience do exercise an influence on judgments of mental activity. However, there is a big explanatory gap between showing that a connection exists and providing a full description of such a relation. One way to attempt closing that gap is to ask participants explicitly what they think about the relation between mind and brain. Doing so is not without challenges: people's responses to explicit questions are often more susceptible to wording artifacts, explicit questioning may trigger ad hoc answers that fail to align with the background beliefs regularly held, and answers may be altogether unreliable if people lack well-established theories. These limitations notwithstanding, listening to people's insights about their beliefs of mind/brain relation may enrich our understanding of such lay theories.

Folk Beliefs About Brain and Mind

Developmental psychology has had a long-standing interest in understanding children's conceptual development. Thus, it comes as no surprise that much of the pioneering work on folk theory of mind and its relation to the brain can be traced back to developmental psychologists (e.g., Lillard, 1998; Johnson & Wellman, 1982). Such research has found that in Western cultures, the mind is often identified

with the brain (Lillard, 1998). When asked “Do you need the brain to ____?”, both adults and elementary school children endorse the view that the brain is necessary for all sorts of human psychological activities. These include emotions such as feeling sad or feeling curious, and senses such as hearing and seeing, but also cognitive acts like thinking and knowing, as well as reading and writing. It includes motor tasks like talking and walking and, in the case of adult participants, even involuntary tasks like coughing and blinking (Johnson & Wellman, 1982). In other words, when asked about the functions of mind and brain, elementary school children and adults alike treat the brain as responsible for the functions of the mind.

Both ninth graders and adults reject that the mind could exist in the absence of the brain, and both of them localize mind and brain in the head. However, ninth graders and adults say that, unlike the brain, the mind is nonmaterial and thus could not be seen nor touched even if the head were opened up. In contrast, young children seem to have a different ontology of the mind. For example, first graders conceive the mind as a material entity that could be seen and touched as much (or as little) as the brain; first graders also tend to construct mind and brain as independent entities, and claim that a mind could exist without a brain (Johnson & Wellman, 1982). Adults and older children also believe that the mind has temporal cohesiveness: they understand that it is the same mind which, encompassing various cognitive processes and states, is being used at different times (Johnson & Wellman, 1982).

Folk Beliefs About Brain and Self

Interestingly, beliefs about spatial and temporal cohesiveness apply not only to the mind but also to the self. The cohesiveness of the self is nicely illustrated by Descartes’ famous inference “I think, therefore I am”. In this statement, Descartes assumes the existence of a self (“I”) doing the thinking. More generally, both adults and children localize the *self* near the eyes (Anglin, 2014; Bertossa, Besa, Ferrari, & Ferri, 2008; Starmans & Bloom, 2012). Besides this conceptualization of the self as the experiencer (the “I”), there is also a conceptualization of the self as an object (the “me”). In that regard, adults think of themselves not as a disparate collection of thoughts and dispositions but rather as a cohesive unit somewhat stable over time, especially as it projects into the future (Moore, Lemmon, & Skene, 2001; Neisser, 1988; Quoidbach, Gilbert, & Wilson, 2013). In other words, the concept is temporally extended to also include the past self and the future self, with interesting asymmetries between the two. For example, adults of all ages think that their values, preferences, and personality traits, having evolved in the past, have now reached a stasis that protects them against further change (Quoidbach et al., 2013). For retrospective judgments, people favor downward comparisons, especially for the distant past. In one such study, college students were given a list of positive attributes, such as willingness to stand up for one’s beliefs, or having good social skills. Students had to indicate the degree to which they possessed each attribute

relative to their same-aged peers, on a scale from 0 (much less than most) to 10 (much more than most). Students assessed the self twice: first as they remembered it at age 16 and then as they knew it at present time; maybe not surprisingly, the rates of positive attributes at age 16 were substantially lower than at their current age. Replications at other ages ruled out an account based on poor adolescent skills (Wilson & Ross, 2001). Instead, the results are best explained by temporal self-appraisal theory, according to which people are motivated to enhance their perception of their current self. In pursuit of this goal, people implicitly make downward comparisons with their former self, as long as the former self is distant enough that it can be plausibly rejected from the current self-conception (Peetz & Wilson, 2008). Self-appraisal theory highlights that the concept of self is not immune to motivated cognition. This should not be a surprise, given that the concept of the self is in itself part of a lay theory informed by semantic knowledge about the mind, as well as by autobiographical memory (Neisser, 1988).

Besides the distinctions with the distant past, there are also distinctions between near and distant future in the conception of the self. For the near future, people adopt a concept of self that is mostly concrete, specific, and context dependent; but for the distant future, people favor instead an abstract self, closest to the *true* or essential self (Wakslak, Nussbaum, Liberman, & Trope, 2008). This latter distinction highlights an important point, namely that the concept of the self is hierarchically organized, with some traits being more central and others being more peripheral (Markus & Wurf, 1987; Sedikides, 1995).

The central self can be defined as the person you *truly* are (i.e., your *true* self) so that if you lacked those attributes you would be a different person; sometimes, the term ‘core self’ is used as synonymous. When asked to describe their ‘central’ self, people often make reference to moral traits, such as “honesty” and “kindness” (Fernandez-Duque & Schwartz, 2016; Goodwin, Piazza, & Rozin, 2014). In other words, when asked to describe who they *truly* are, what people volunteer are uniquely human traits, traits that Descartes believed did not belong in the brain. And while philosophers have long ago moved beyond Cartesian dualism, it remains a legitimate scientific contention that ordinary folk still hold to this belief (Bloom, 2004). According to this hypothesis, people would be willing to admit that the brain is the substrate of cognitive functions and many psychological traits, but would reserve a special nonmaterial place for traits that define who they truly are.

To test this hypothesis, we asked a group of people residing in the USA (recruited through Amazon’s Mturk) to judge whether the brain was “more responsible for the CORE attributes of your self or for the PERIPHERAL attributes of your self” using a 100-point bipolar scale (Fernandez-Duque & Schwartz, 2016). We explained to the 172 participants that “the core self is who you truly are [...] so that if you lacked those attributes you would be a different person” while the peripheral self included “things that describe you but don’t define you [...] so that if you didn’t have those attributes, you would still be the same person.” Contrary to the hypothesis, participants embraced the brain as the underlying substrate of their central self, that is, of who they *truly* are (Fernandez-Duque & Schwartz, 2016). We also asked another 210 participants about the neuroscience contribution toward 18

different personality traits. Once again, the brain contribution for traits closer to the central self was deemed larger than for more peripheral traits. As expected, there was quite a bit of variability among traits, with perceived contributions of the brain ranging from 54% (for laziness) to 91% (for intelligence). In the future, studies probing a larger number of traits will help identify the trait attributes (e.g., volitional control, desirability, etc.) that best predict the perceived brain contribution.

This belief that the brain is the underlying substrate to people's *true* self and personality is also apparent in caregivers' reports of frontotemporal dementia patients. Brains affected by frontotemporal dementia are lesioned in areas similar to Phineas Gage's, and as a consequence these patients with frontotemporal dementia often exhibit similar behavioral and moral transgressions (Fernandez-Duque & Black, 2007; Fernandez-Duque, Hodges, Baird, & Black, 2010). Friends and families often report that the personal identity of the patient has changed since the start of the disease and that the patient "seems like a stranger" and "is not the same person underneath." At an intuitive level, these caregivers are endorsing the belief that changes to their loved ones' *true* self was brought about by pathological changes in their brains (Strohming & Nichols, 2015). As such, it is an example of what in the folk psychiatry literature has been called neuro-essentialism, which is "the belief that brains and their abnormalities define and determine identity" (Haslam, 2011).

The Possibility of Dualism

Many religious beliefs that are popular across the world depend on a dualist concept of mind/brain relation (or at the very least of the soul/brain relation). For example, beliefs about the afterlife require the existence of a nonmaterial substance separate from the body (Greely & Hout, 1999). Furthermore, there seems to be a clear positive correlation between popular dualism and other beliefs that seem dependent on it, such as beliefs in the afterlife, paranormal beliefs, and some religious beliefs (Fernandez-Duque & Schwartz, 2016; Riekkki, Lindeman, & Lipsanen, 2013). Based on the evidence like this, as well as some of the developmental literature, some researchers have argued that children start as dualists, and become materialists only years later—if at all—through formal education (Bloom, 2004). According to this view, people learn in school, and through the internet and other media, that "the brain underlies the mind" the same way that people learn all sorts of strange, unintuitive scientific facts (Bloom, 2004). The evidence for and against dualism stems from various fronts and is described in detail in other chapters of this book (Haslam, Chapter "The Origins of Lay Theories: The Case of Essentialist Beliefs"; Forstmann & Burgmer, Chapter "Antecedents, Manifestations, and Consequences of Belief in Mind–Body Dualism").

However, in interpreting these data, and in interpreting data on folk theories of mind/brain more generally, it is important to keep in mind that any characterization of common sense beliefs about the mind/brain relation needs to contend with the fact that the concept of the 'mind' is not a monolithic construct but rather a multifaceted

one; therefore it is possible—and even likely—that common sense beliefs about the mind may similarly include a constellation of different beliefs, with some psychological states deemed more brain based than others. For example, American adults tend to cluster mental states into two dimensions, an experiencing/feeling dimension that includes psychological states such as the feelings of hunger, fear, and joy and an agency/cognitive dimension that includes psychological states such as self-control, morality and memory (Gray, Gray, & Wegner, 2007). A biological brain appears necessary for experiencing things, such as hunger, joy, or pleasure, as people do not attribute those experiences to God or a robot. In contrast, for agency, a biological brain seems neither necessary nor sufficient: God and robots are deemed high on agency while newborns and frogs are denied it (Gray, Gray, & Wegner, 2007). Similarly, people resist attributing experiences to brainless corporations (“Acme is feeling pain”) but accept attributions of agency to them (“Acme Corp plans to change its corporate image” Knobe & Prinz, 2008).

Why Is It Imperative to Understand Lay Theories of Mind/Brain?

The issues discussed in this chapter are important to understand not only because they enrich our description of how humans categorize and conceptualize the world—in the tradition of past research on folk physics (McCloskey & Kohl, 1983), folk biology (Carey, 1985), and folk mentalizing (Wellman, Cross, & Watson, 2001)—but maybe more importantly because the decisions humans make, and the social worlds that they construct, derive directly from the lay theories and beliefs they hold regarding those worlds. Thus, different conceptualizations of the mind/brain relation should have profound implications for public policy in mental health research and practice. By better understanding those conceptualizations we might be able to modify them, and in doing so, we may be able to modify our destiny. These aspirations sound lofty and vague, so some concrete elaboration is in order.

Let us start by stating the obvious: the natural world does not care about the theories humans create to explain it. Alchemy in the seventeenth century may have proposed a theory to turn copper into gold, but no amount of theorizing was ever going to make that happen. By contrast, the social world is quite susceptible to the theorizing we humans do in order to explain it. The distinguished psychologist Barry Schwartz has illustrated this point in his explanation of why we work (Schwartz, 2015). According to Adam Smith, the father of free market economic theory, we work for pay, nothing more and nothing less. A workplace in which workers’ only motivation is thought to be monetary is likely to be designed lacking any other sources of motivation; after all, why waste resources promoting workers’ sense of accomplishment, or creating a supportive social life in the workplace, if we *know* that workers only care about their paycheck? In such a devoid environment, workers would not find any reason to work other than their salary, and when asked, they will confirm our initial theory.

If this analysis is correct, then we *discover* the natural world but we *create* the social world. When it comes to lay theories of the mind/brain relationship, that social world includes, among other things, the treatment of mental disorders, the funding priorities for mind/brain research, and the implementation of our legal and educational systems.

Consider, for example, the treatment of mental disorders. When mental disorders, such as ADHD or generalized anxiety disorder are explained by appeal to biological information, people become overly pessimistic about their prognosis (Lebowitz & Ahn, 2014; Lebowitz, Pyun, & Ahn, 2014). When therapists hear of mental diseases such as depression and OCD in biological terms, they become less sympathetic toward the patients (Lebowitz & Ahn, 2014). Therapists' beliefs are quite malleable, so that those with a medical degree are more inclined to think of the disorders as medically based (Kim, Ahn, Johnson, & Knobe, 2016). More arbitrary biases are present too; when the disorder is described abstractly in terms of symptoms, therapists think of it as biologically based, but when described concretely in relation to an individual patient, the same therapists become more inclined to think of the disorder as psychologically based, and less susceptible to medical treatment; this is true even for those therapists who are medically trained (Lebowitz, Rosenthal, & Ahn, 2016). Therapists and laypersons conceptualize mental disorders along a single continuum that spans from disorders considered highly biological (e.g., autism) to disorders considered highly psychosocial (e.g., adjustment disorders), thus ignoring the quite likely scenario of dual contributions from biological and social factors (Ahn, Proctor, & Flanagan, 2009). The goal of these examples is not to dwell on this very interesting literature (for a deep analysis, see Furnham, chapter "[How Lay Theories Influence our Mental Health](#)") but rather to illustrate the claim that our folk theories of the mind/brain relation have a profound impact on how we approach and try to solve the problems in front of us, both in terms of clinical practice and of public policy.

Another illustrative example comes from the funding priorities for mind/brain research at the United States' National Institute of Mental Health, which in the last decade, under the directorship of Tom Insel, a neuroscientist known for his work on hormonal control of monogamy in mammals, has redirected its focus away from social science and toward neuroscience, where it is now almost exclusively focused (Markovitz, 2016). It seems reasonable to speculate that the folk theory at the helm regarding the relation between mind and brain has been at least partly responsible for those changes. It also seems reasonable to ask whether those changes in priorities reflect the values of the citizenry, as expressed by folk beliefs. For example, should the taxonomy of mental disorders be organized by identifying symptom clusters, or should it instead be built bottom up from genes and neurobiology? Should promising behavioral therapies receive funding for effectiveness testing, or lose such funding due to the lack of a neurological correlate in their proposed mechanism? Should the effectiveness of potential treatments for Alzheimer's disease be judged based on their ability to remove the biological substrate of the disease (i.e., presumably plaques) or by their ability to improve behavior (e.g., episodic memory)? As the preceding discussion makes clear, there is a lack of

consensus among scientists about the level of analysis at which mental disorders ought to be conceptualized and treatment delivered; the answer to such questions depends to a great extent on the particular disease under discussion, and often the most effective treatments combine interventions at both neural and psychological levels (i.e., drugs and talk therapy).

In the field of education, the situation is quite different. Although there has been a fair amount of hype surrounding neuroscience and education for the last 20 years, the level of analysis at which educational gains are maximized is undoubtedly psychological rather than neurological (Bowers, 2016). This should not be surprising; after all, the primary outcomes of education are behavioral: we want children to learn to read, do math, develop critical thinking skills, and so forth. And unlike mental disorders, the treatment options are exclusively behavioral: in order to foster children's phonemic decoding, we sound out letters, in order to foster a number sense, we draw a number line, and so forth. At most, the potential of neuroscience to affect education is likely limited to low-level behaviors, such as reading, rather than more complex behaviors such as collaborating with a classmate or writing an essay. Nonetheless, how likely are scientific theories of mind/brain relation to influence education in years to come? To start answering this question, it is helpful to start with a brief history of the mind/brain relation in science.

In the early 1800s, phrenologists had aimed to divide the mind into its constituent mental faculties but had failed spectacularly, due to a lack of empirical rigor and a penchant for ill-conceived categorization. By the late 1800s, Paul Broca had overcome some of these limitations by discovering that speech could be mapped to a specific region of the brain (Broca, 1861; Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007). For many decades afterwards, the prevalent paradigm remained trying to relate large complex task capacities—speech, memory, motor control—to similarly large brain regions (Thiebaut de Schotten et al., 2015).

But in the '70s, cognitive psychologists started to break down those large cognitive capacities into smaller mental operations. To achieve this, they administered relatively simple tasks, contrasted nearly identical experimental conditions, and measured response times with millisecond precision. This way, cognitive psychologists were able to isolate what they referred to as “elementary mental operations,” the building blocks from which complex cognitive tasks are made (Posner, 1978). In the following decades, proponents of these “chronometric explorations of the mind” would convincingly argue that such elementary mental operations constituted the appropriate level of analysis at which to map mind and brain (Posner & Raichle, 1994). The idea became the main tenet of the new field of cognitive neuroscience, and helped cognitive neuroscience move past not only phrenology's but also Paul Broca's conception of the mind.

This scientific conceptualization of the mind/brain relation has provided cognitive neuroscientists with the necessary theoretical models to uncover the neural bases of reading (Dehaene, 2009), mental calculation (Dehaene, 2011), and working memory, as well as many other cognitive processes, each with its own set of elemental mental operations. As a consequence, sophisticated neural models of dyslexia and dyscalculia have been developed over the last 20 years,

and neuroimaging studies can now predict above chance which children will develop dyslexia, and which of them will benefit from treatment, thus opening the door to personalized educational treatment (Butterworth, Varma, & Laurillard, 2011; Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015). Equally important, the models have had an impact on educational policy: based on the neuroscience evidence, the National Institute of Child Health and Human Development now defines dyslexia as “a specific learning disability that is neurobiological in origin.” As this example nicely illustrates, the scientific theory of the mind/brain relation has already started to influence some aspects of educational policy and practice, although the extent of such influence in the future remains to be seen.

In this section, I have argued that a better understanding of current theories of the mind/brain relation is important, not only as a basic scientific endeavor, but also as a tool for public policy and practice. The impact of such theories is found in areas as diverse as mental health, education, and science funding, as well as in the field of psychology and law. Unresolved issues remain, such as the potential gap between the folk theory and the expert theory, and how to adjudicate in cases in which folk theory and expert theory disagree.

Summary

In this chapter, I have tried to answer a deceptively simple question: Why is neuroscience so alluring? But to ask this question is to ask about current lay theories of the mind/brain relation. I started by reviewing the allure of neuroscience explanations for psychological phenomena. I showed that although neuroscience is a prestigious science, this does not explain its allure. Nor can the allure of neuroscience explanations be explained by the neuroimages that often accompany them. Instead, the allure stems from a folk theory of the mind that conceptualizes the brain as the engine of the mind. Neuroscience is alluring for explaining psychological phenomena because of its reductive appeal.

Next, I reviewed the effect of neuroscience on judgments of responsibility and free will. This is another way to ask about the folk theory of the mind/brain relation: how the brain contributes to uniquely human attributes such as free will and morality. I showed that although ordinary folk shy away from free will and moral condemnation when actions are couched in terms of brain function, the reason for this is not that people think of free will as something that cannot exist in a fully deterministic material world. Rather, people conceptualize free will as the ability to make choices consistent with one’s desires, reasonably free of constraints. To the extent that a neurological disorder disrupts these underlying psychological components of ‘free will’, people attribute a reduction of ‘free will’. Thus, we excuse the moral transgressions of patients who, like Phineas Gage, have abrupt brain lesions not because our knowing of his brain lesion turns us into hard determinists skeptical of metaphysical free will, but rather because our knowing of his lesion reminds us of Gage’s reduced mental capacities to make choices consistent with his

desires and free of unreasonable constraints. In contrast, the existence of a normally functioning brain ought not pose a challenge to free will because a normal brain—by definition—has a correspondingly normal psychology.

The next stop on our tour of uniquely human capacities and their relation to the brain was the notion of the self. Do ordinary folk believe that their brains are who they *truly* are? Recent evidence suggests that they do, at least for Western educated cultures. The chapter also briefly discussed mind–body dualism and neuro-essentialism, topics that get the attention they deserve in other chapters of the book (Haslam, Chapter “[The Origins of Lay Theories: The Case of Essentialist Beliefs](#)”; Forstmann & Burgmer, Chapter “[Antecedents, Manifestations, and Consequences of Belief in Mind–Body Dualism](#)”).

Finally, I concluded by exploring possible policy implications of the allure of neuroscience and current theories of the brain/mind relation. I have argued that the impact of such theories is found in areas as diverse as mental health, education, and science funding. As such, I hope they help illustrate the profound implications that differences in the conceptualizations of the mind/brain relation may have, not only for our understanding of human cognition but perhaps more importantly, for human society.

References

- Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians’ beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, 33(2), 147–182. doi:[10.1111/j.1551-6709.2009.01008.x](#)
- Anglin, S. M. (2014). I think, therefore I am? Examining conceptions of the self, soul, and mind. *Consciousness and Cognition*, 29, 105–116. doi:[10.1016/j.concog.2014.08.014](#)
- Aspinwall, L. G., Brown, T. R., & Tabery, J. (2012). The double-edged sword: does biomechanism increase or decrease judges’ sentencing of psychopaths? *Science*, 337(6096), 846–849. doi:[10.1126/science.1219569](#)
- Beall, E., Meindl, P., Iyer, R., Monterosso, J., Graham, J., & Schwarz, B. (2013). *Biology or bad character? Brain scan images increase determinism and decrease moral attributions for criminal acts*. Paper presented at the Annual Conference of the Society for Personality and Social Psychology, New Orleans, LA.
- Bering, J. M., & Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, 40(2), 217–233. doi:[10.1037/0012-1649.40.2.217](#)
- Bertossa, F., Besa, M., Ferrari, R., & Ferri, F. (2008). Point zero: A phenomenological inquiry into the seat of consciousness. *Perceptual and Motor Skills*, 107(2), 323–335.
- Bloom, Paul. (2004). *Descartes’ baby: How the science of child development explains what makes us human*. New York: Basic Books.
- Bowers, J. S. (2016). The practical and principled problems with educational neuroscience. *Psychological Review*, 123(5), 600–612. doi:[10.1037/rev0000025](#)
- Broca, P. (1861). Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech). *Remarques sur le siège de la faculté du langage articulé, suivies d’une observation d’aphémie (perte de la parole)*. *Bulletin de la Société Anatomique*, 6, 330–357.
- Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science*, 332(6033), 1049–1053. doi:[10.1126/science.1201536](#)
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, Mass.: MIT Press.

- Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. New York: Viking.
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed.). New York: Oxford University Press.
- Demertzi, A., Liew, C., Ledoux, D., Bruno, M. A., Sharpe, M., Laureys, S., et al. (2009). Dualism persists in the science of mind. *Annals of the New York Academy of Science*, 1157, 1–9. doi:10.1111/j.1749-6632.2008.04117.x
- Descartes, R. (1641) (1984/1641). Meditations on first philosophy. In (Cottingham, Trans.) R. J. Stoothoff & D. Murdoch (Eds.), *The philosophical writings of René Descartes* (Vol. 2, pp. 1–62). Cambridge: Cambridge University Press.
- Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130(Pt 5), 1432–1441. doi:10.1093/brain/awm042
- Eriksson, K. (2012). The nonsense math effect. *Judgment and Decision Making*, 7(6), 746–749.
- Fernandez-Duque, D., & Black, S. E. (2007). Metacognitive judgment and denial of deficit: Evidence from frontotemporal dementia. *Judgment and Decision Making*, 2(6), 359–370.
- Fernandez-Duque, D., Evans, J., Christian, C., & Hodges, S. D. (2015). Superfluous neuroscience information makes explanations of psychological phenomena more appealing. *Journal of Cognitive Neuroscience*, 27(5), 926–944. doi:10.1162/jocn_a_00750
- Fernandez-Duque, D., Hodges, S. D., Baird, J. A., & Black, S. E. (2010). Empathy in frontotemporal dementia and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, 32(3), 289–298. doi:10.1080/13803390903002191
- Fernandez-Duque, D., & Schwartz, B. (2016). Common sense beliefs about the central self, moral character, and the brain. *Frontiers in Psychology*, 6. doi:10.3389/fpsyg.2015.02007
- Gabrieli, J. D., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*, 85(1), 11–26. doi:10.1016/j.neuron.2014.10.047
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148–168. doi:10.1037/a0034726
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315 (5812), 619. doi:10.1126/science.1134475
- Greely, A. M., & Hout, M. (1999). Americans' increasing belief in life after death: Religious competition and acculturation. *American Sociological Review*, 64, 813–835.
- Greene, E., & Cahill, B. S. (2012). Effects of neuroimaging evidence on mock juror decision making. *Behavioral Sciences and the Law*, 30(3), 280–296. doi:10.1002/bsl.1993
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 359 (1451), 1775–1785. doi:10.1098/rstb.2004.1546
- Gurley, J. R., & Marcus, D. K. (2008). The effects of neuroimaging and brain injury on insanity defenses. *Behavioral Sciences and the Law*, 26(1), 85–97. doi:10.1002/bsl.797
- Harlow, J. (1868). Recovery from the passage of an iron bar through the head. *Massachusetts Medical Society*, 2, 327–347.
- Haslam, N. (2011). Genetic essentialism, neuroessentialism, and stigma: commentary on Dar-Nimrod and Heine (2011). *Psychological Bulletin*, 137(5), 819–824. doi:10.1037/a0022386
- Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition*, 155, 67–76. doi:10.1016/j.cognition.2016.06.011
- Johnson, C. N., & Wellman, H. M. (1982). Children's developing conceptions of the mind and brain. *Child Development*, 53(1), 222–234.
- Keil, F. C., Lockhart, K. L., & Schlegel, E. (2010). A bump on a bump? Emerging intuitions concerning the relative difficulty of the sciences. *Journal of Experimental Psychology: General*, 139(1), 1–15. doi:10.1037/a0018319

- Kim, N. S., Ahn, W. K., Johnson, S. G., & Knobe, J. (2016). The influence of framing on clinicians' judgments of the biological basis of behaviors. *Journal of Experimental Psychology: Applied*, 22(1), 39–47. doi:10.1037/xap0000070
- Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences*, 7, 67–83.
- Lebowitz, M. S., & Ahn, W. K. (2014). Effects of biological explanations for mental disorders on clinicians' empathy. *Proceeding National Academy of Science U S A*, 111(50), 17786–17790. doi:10.1073/pnas.1414058111
- Lebowitz, M. S., Pyun, J. J., & Ahn, W. K. (2014). Biological explanations of generalized anxiety disorder: Effects on beliefs about prognosis and responsibility. *Psychiatry Services*, 65(4), 498–503. doi:10.1176/appi.ps.201300011
- Lebowitz, M. S., Rosenthal, J. E., & Ahn, W. K. (2016). Effects of biological versus psychosocial explanations on stigmatization of children with ADHD. *Journal of Attentional Disorders*, 20(3), 240–250. doi:10.1177/1087054712469255
- Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin*, 123(1), 3–32.
- Markovitz, J. C. (2016). There is such a thing as too much neuroscience. *New York Times*, October 14. Retrieved from <http://www.nytimes.com/2016/10/15/opinion/theres-such-a-thing-as-too-much-neuroscience.html>
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38, 299–337.
- McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, 107(1), 343–352. doi:10.1016/j.cognition.2007.07.017.
- McCloskey, M., & Kohl, D. (1983). Naive physics: The curvilinear impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(1), 146–156.
- Michael, R. B., Newman, E. J., Vuorre, M., Cumming, G., & Garry, M. (2013). On the (non) persuasive power of a brain image. *Psychonomic Bulletin and Review*, 20(4), 720–725. doi:10.3758/s13423-013-0391-6
- Monroe, A. E., Brady, G., & Malle, B. F. (2016). This isn't the free will worth looking for: General free will beliefs do not influence moral judgments, agent-specific choice ascriptions do. *Social Psychological and Personality Science*, 27, 100–108. doi:10.1177/1948550616667616
- Monroe, A. E., & Malle, B. F. (2010). From uncaused will to conscious choice: The need to study, not speculate about, people's folk concept of free will. *Review of Philosophy and Psychology*, 1, 211–224.
- Monterosso, J., Royzman, E. B., & Schwartz, B. (2005). Explaining away responsibility: Effects of scientific explanation on perceived culpability. *Ethics and Behavior*, 15(2), 139–158. doi:10.1207/s15327019eb1502_4
- Moore, C., Lemmon, K., & Skene, K. (Eds.). (2001). *The self in time: Developmental Perspectives*. Mahwah, NJ: Erlbaum.
- Nahmias, E. (2006). Folk fears about freedom and responsibility: Determinism vs. reductionism. *Journal of Cognition and Culture*, 6(1–2), 215–237.
- Nahmias, E. (2016). Free will as a psychological accomplishment. In D. Schmidtz & C. Pavel (Eds.), *The Oxford handbook of freedom*: OUP.
- Nahmias, E. (2017). Your brain as the source of free will worth wanting: Understanding free will in the age of neuroscience. In G. Caruso & O. Flanagan (Eds.), *Neuroexistentialism: Meaning, morals, and purpose in the age of neuroscience*. Oxford: Oxford University Press.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, 18, 561–584.
- Nahmias, E., Shepard, J., & Reuter, S. (2014). It's OK if 'my brain made me do it': People's intuitions about free will and neuroscientific prediction. *Cognition*, 133(2), 502–516. doi:10.1016/j.cognition.2014.07.009

- Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology*, *1*(1), 35–59.
- Nichols, S. (2011). Experimental philosophy and the problem of free will. *Science*, *331*(6023), 1401–1403. doi:[10.1126/science.1192931](https://doi.org/10.1126/science.1192931)
- Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs*, *41*(4), 663–685.
- Peez, J., & Wilson, A. E. (2008). The temporally extended self: The relation of past and future selves to current identity, motivation, and goal pursuit. *Social and Personality Psychology Compass*, *2*, 2090–2106.
- Posner, M. I. (1978). *Chronometric explorations of mind: The third Paul M. Fitts lectures, delivered at the University of Michigan, September 1976*. Hillsdale, N.J.: L. Erlbaum Associates
- Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York: Scientific American Library.
- Quoidbach, J., Gilbert, D. T., & Wilson, T. D. (2013). The end of history illusion. *Science*, *339*(6115), 96–98. doi:[10.1126/science.1229294](https://doi.org/10.1126/science.1229294)
- Racine, E., Bar-Ilan, O., & Illes, J. (2005). fMRI in the public eye. *Nature Review Neuroscience*, *6*(2), 159–164. doi:[10.1038/nrn1609](https://doi.org/10.1038/nrn1609)
- Riecki, T., Lindeman, M., & Lipsanen, J. (2013). Conceptions about the mind-body problem and their relations to afterlife beliefs, paranormal beliefs, religiosity, and ontological confusions. *Advances in Cognitive Psychology*, *9*(3), 112–120. doi:[10.2478/v10053-008-0138-5](https://doi.org/10.2478/v10053-008-0138-5)
- Roskies, A. (2006). Neuroscientific challenges to free will and responsibility. *Trends Cogn Sci*, *10*(9), 419–423. doi:[10.1016/j.tics.2006.07.011](https://doi.org/10.1016/j.tics.2006.07.011)
- Schwartz, B. (2015). *Why we work* (First TED books hardcover edition September 2015. ed.).
- Schweitzer, N. J., & Saks, M. J. (2011). Neuroimage evidence and the insanity defense. *Behavioral Sciences and the Law*, *29*(4), 592–607. doi:[10.1002/bsl.995](https://doi.org/10.1002/bsl.995)
- Sedikides, C. (1995). Central and peripheral self-conceptions are differentially influenced by mood: tests of the differential sensitivity hypothesis. *Journal of Personality and Social Psychology*, *69*(4), 759–777.
- Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J., Clark, C. J., Schooler, J. W., et al. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science*, *25*(8), 1563–1570. doi:[10.1177/0956797614534693](https://doi.org/10.1177/0956797614534693)
- Sloman, S. A., & Rabb, N. (2016). Your understanding is my understanding: Evidence for a community of knowledge. *Psychological Science*, *27*(11), 1451–1460. doi:[10.1177/09567976166662271](https://doi.org/10.1177/09567976166662271)
- Stanovich, K. E. (1989). Implicit philosophies of mind—The dualism scale and its relation to religiosity and belief in extrasensory perception. *Journal of Psychology*, *123*(1), 5–23.
- Starmans, C., & Bloom, P. (2012). Windows to the soul: Children and adults see the eyes as the location of the self. *Cognition*, *123*(2), 313–318. doi:[10.1016/j.cognition.2012.02.002](https://doi.org/10.1016/j.cognition.2012.02.002)
- Strohming, N., & Nichols, S. (2015). Neurodegeneration and Identity. *Psychological Science*, *26*(9), 1469–1479. doi:[10.1177/0956797615592381](https://doi.org/10.1177/0956797615592381)
- Tal, A., & Wansink, B. (2014). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*. doi:[10.1177/0963662514549688](https://doi.org/10.1177/0963662514549688)
- Thiebaut de Schotten, M., Dell’Acqua, F., Ratiu, P., Leslie, A., Howells, H., Cabanis, E., ... Catani, M. (2015). From Phineas Gage and Monsieur Leborgne to H.M.: Revisiting disconnection syndromes. *Cerebral Cortex*, *25*(12), 4812–4827. doi:[10.1093/cercor/bhv173](https://doi.org/10.1093/cercor/bhv173)
- Wakslak, C. J., Nussbaum, S., Liberman, N., & Trope, Y. (2008). Representations of the self in the near and distant future. *Journal of Personality and Social Psychology*, *95*(4), 757–773. doi:[10.1037/a0012939](https://doi.org/10.1037/a0012939)
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, *20*(3), 470–477. doi:[10.1162/jocn.2008.20040](https://doi.org/10.1162/jocn.2008.20040)
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655–684.
- Wilson, A. E., & Ross, M. (2001). From chump to champ: People’s appraisals of their earlier and present selves. *Journal of Personality and Social Psychology*, *80*(4), 572–584.

Causes and Consequences of the Belief in Free Will

Davide Rigoni, Axel Cleeremans and Marcel Brass

Free will is an ancient problem. Although the existence of free will has been challenged both at the philosophical and scientific levels (e.g., Crick, 1994; Harris, 2012), most laypeople believe they and others have free will. This belief reflects the lay theory that people have the capacity to decide freely—that is, free from internal and external constraints—and are thus responsible for their own actions (Monroe, Dillon, & Malle, 2014; Nahmias, Coates, & Kvaran, 2007). As any other lay theory, the function of the belief in free will is to interpret, predict, and make sense of people's behavior. Regardless of whether free will exists or not, the belief in free will is a cornerstone of our social and cultural life. For instance, most legal systems are based on the lay theory that individuals can act on the basis of their own free will, and can therefore be judged responsible and punishable for their own actions (Rigoni, Sammiceli, & Sartori, 2015).

The fact that most people believe in free will raises a number of questions: why is it the case? Does it actually matter whether people believe in free will? And how does believing in free will affect people's behavior?

In this chapter, we will first describe recent empirical research examining the cognitive and social determinants of the belief in free will. On the one hand, cognitive

D. Rigoni (✉) · M. Brass
Department of Experimental Psychology, Ghent University,
Henri Dunantlaan 2, 9000 Ghent, Belgium
e-mail: davide.rigoni@ugent.be

A. Cleeremans
Consciousness, Cognition and Computation Group, Brussels, Belgium

A. Cleeremans
Center for Research in Cognition and Neurosciences, Brussels, Belgium

A. Cleeremans
ULB Neurosciences Institute, Université Libre de Bruxelles, Brussels, Belgium

research has shown that the belief in free will is grounded in our first-hand, embodied experience of being the agents of our own actions. On the other hand, empirical findings in social psychology have suggested that belief in free will plays a key societal function, namely that it disciplines social and cultural life.

Next, we delineate a recent line of research that focuses on the social, cognitive, and neural consequences of (dis)believing in free will. This research suggests that challenging people's belief in free will can influence social behavior as well as more basic cognitive and neural mechanisms of self-regulation.

In the last section, we briefly outline some potential mechanisms of how abstract beliefs about free will can impact behavior and cognition. Central to the proposal is the idea that beliefs about free will can be thought of in terms of metacognitive judgments about oneself and others. As a form of metacognition, abstract beliefs about free will impact on cognition and on behavior by redescribing and shaping first-order sensory-motor representations.

Cognitive and Social Determinants of the Belief in Free Will

Belief in Free Will and Agentive Control

Empirical research on the folk concept of free will indicates that people's understanding of free will reflects the psychological ability to make a choice in line with one's desires or goals and free of internal (e.g., genetic makeup, personality, mental disorder) or external (e.g., society, nature, God) constraints (Monroe et al., 2014; Monroe & Malle, 2010). Free will can therefore be defined as the capacity to choose freely among different available courses of actions. Although there are individual differences in the extent with which people believe in the existence of free will in an abstract sense, our conscious experience provides us with the clear sense that we can intentionally control our behavior. Our subjective experience of being free is that desires and intentions precede and motivate our actions, and that we are constantly called to make choices about what path to take. Acting on the basis of one's own intentions indeed seems to be a core feature of the folk notion of free will. Reflexes (i.e., involuntary muscular contractions), for example, are not considered acts of free will because they do not imply the formation of an intention. Conversely, when one pushes a button with the intention to get a refreshing drink from a vending machine, one feels in control of and ultimately responsible for this action. In other words, the concept of free will only makes sense for actions that involve some sense of intentional control over one's own decisions and actions.

In the psychological research literature, the subjective feeling of being in control of one's own behavior is typically referred to as the *sense of agency*, which can be specifically defined as the feeling of control over self-produced actions (Gallagher, 2012; Moore & Obhi, 2012). Going beyond the common intuition that free will requires the sense of agency, Aarts and van den Bos directly tested the hypothesis

that the strength of individuals' beliefs in free will depends on the extent to which they experience themselves as the *agents* of self-produced actions (Aarts & van den Bos, 2011). In this study, sense of agency was measured through the intentional binding paradigm, in which participants are asked to produce a tone by pressing a key and then provide a time estimation of both the action and the tone (i.e., the action effect). It was previously demonstrated that the action and its effect are perceived as closer to each other in time when the action is produced intentionally by the participant, as compared to when the movement is performed passively and unintentionally (Haggard, Clark, & Kalogeras, 2002). This phenomenon, referred to as *intentional binding*, is widely interpreted as a cognitive marker of the sense of agency (Moore & Obhi, 2012). In two experiments, Aarts and van den Bos (2011) provided empirical evidence that individuals' explicit belief in free will, as measured through self-report questionnaires (e.g., "*People have complete control over the decisions they make*"; Paulhus & Carey, 2011), is predictive of the strength of intentional binding (Experiment 1), even when action outcomes are primed unconsciously (Experiment 2). These findings indicate that the strength of people's high level and conceptual belief in free will depends, to a certain extent, on the subjective sense of control over self-produced actions: The more people *feel* they are in control of their actions, the stronger their explicit belief in free will.

Further evidence that subjective sense of control influences people's belief in free stems from a series of studies that tested the hypothesis that experiencing either a chronic or a temporary lack of intentional control over one's own bodily reactions can lower people's belief in free will (Ent & Baumeister, 2014). In one study, the authors measured belief in free will in people with either epilepsy or panic disorders, two medical conditions characterized by a lack of control over one's own actions and behavior. They found that both clinical conditions were associated to weaker belief in free will as compared to control participants. Although these observations are correlational, it can be assumed that suffering from a medical disorder that induces a lack control over behavior lowers one's belief in free will, rather than the other way around. Interestingly, the authors also found that temporary changes in the feeling of control over one's own body, such as experiences of tiredness, sexual desire, or hunger can affect people's abstract belief in free will (Ent & Baumeister, 2014): the more intense the reported physiological need, the lower the belief in free will. These results indicate that experiencing limited intentional control over bodily responses, either chronically or temporarily, can reduce people's belief in free will.

Additional support for the hypothesis that 'first-hand' experience of agentic control over behavior is a predictor of abstract beliefs about free will comes from a series of studies showing that people do not attribute the same degree of free will to themselves as to others. Pronin and Kugler (2010) provided evidence that people believe their behavior is driven by intentions and free choices to a larger extent as compared to others—i.e., they attribute more free will to themselves than to others. While these effects can also involve other cognitive biases (e.g., actor–observer bias; Jones & Nisbett, 1972), we can speculate that the subjective experience of

control over one's own actions and behavior, which is obviously absent for others' behavior, reinforces people's belief in their own free will.

In sum, there is empirical evidence that the strength of one's belief in free will depends, at least in part, on the subjective experience of being the agent of one's own actions and behavior. Put differently, the abstract belief that we have free will is grounded in the low-level mechanisms that underpin our sense of agency. This 'sensory-motor' account of free will could explain why the belief in free will is a cultural universal. Although different cultures may differ in their conceptualizations of free will and related ideas, such as moral responsibility (Miller & Bersoff, 1992) and independent agency (Kashima et al., 1995), the belief that humans have free will capacities seems to be virtually present in all societies (Sarkissian et al., 2010).

Belief in Free Will and Need for Moral Rules

While the studies mentioned above highlight the sensory-motor foundations of the belief in free will, other lines of research have examined its social basis. It has been proposed that the concept of free will has been acquired through evolutionary processes in order for individuals to live harmoniously in a social and cultural group (Baumeister, 2008; Dennett, 2003). This functional interpretation of the belief in free will is analogous to hypotheses about other types of beliefs, such as the belief in God (Laurin, 2017). Here, the core idea is that the belief in free will is necessary to hold people morally responsible for their own actions. In other words, people's belief in free will would have emerged as a result of extended social interactions and a shared culture that promotes socially desirable behavior.

This hypothesis has received empirical support from both correlational and experimental research on the relationship between the belief in free will and social variables related to moral judgments and punitive attitudes. For instance, people's belief in free will is associated with conservative attitudes (e.g., authoritarianism), the belief in a 'just' world (i.e., the belief that people are responsible for their behavior and must accept the consequences), and with punitive attitudes toward wrongdoers (Carey & Paulhus, 2013; Crescioni, Baumeister, Ainsworth, Ent, & Lambert, 2015). These correlations indicate that believing in free will is related to worldview perspectives emphasizing compliance with moral rules and societal norms, thereby supporting the idea that the belief in free will plays a key societal function.

Empirical support for the proposal that the belief in free will is functional for holding others responsible for their behavior can be found in a series of studies in which the moral valence of an observed behavior was manipulated experimentally. It has been shown that wrongful behaviors have a stronger impact on judgments of responsibility and intentional control, as compared to morally good or neutral behaviors (Knobe & Fraser, 2008; Young & Phillips, 2011). In other words, the attribution of free will capacities may be a consequence, rather than the starting

point, of harmful and immoral behavior. In a series of studies, Clark and colleagues directly tested the hypothesis that individuals' belief in free will is modulated by the motivation to hold people responsible for their wrongful actions (e.g., Clark et al., 2014). In one of their studies, participants were asked to either read a newspaper article reporting an immoral behavior (e.g., a corrupt judge) or a control article about a job search. After reading the article, participants were asked to report their belief in free will on a standardized scale (Paulhus & Carey, 2011). The rationale behind this manipulation was that considering an immoral act should lead people to perceive more intentionality and more controllability in the wrongdoer and should prompt the need to punish them. This, in turn, should increase their belief in free will. Results showed that participants reading the newspaper article about the immoral behavior reported stronger belief in free will as compared to those who read the control text. The authors expanded this observation in follow-up studies that also involved more realistic contexts (e.g., a student cheating in the classroom) as well as nation-level survey data. These findings have consistently demonstrated that individuals' belief in free will—as well as their desire for punishment—is strengthened when they are confronted with unlawful behavior. Importantly, people not only attribute more free will to the wrongdoers, but they also report stronger belief in free will for people in general, suggesting that the necessity to attribute free will to others is not contingent on the current situation, but rather involves a pervasive psychological process.

The Consequences of Believing or Disbelieving in Free Will

The research outlined above is focused on the cognitive and social sources of individuals' belief in free will. Other researchers have examined the question of whether the belief in free will can be manipulated, and whether such changes can in turn affect cognition and social behavior. As we describe in the following section, these studies show that top-down changes in the belief in free will has effects at all levels, from moral behavior to basic neural and cognitive mechanisms underlying self-regulation.

Challenging Free Will Affects Social Behavior

Most people believe in free will (Baumeister, Masicampo, & Dewall, 2009; Nahmias, Morris, Nadelhoffer, & Turner, 2005), but it has been repeatedly shown that this belief can be reduced by exposing people to philosophical and/or scientific views that deny free will. For instance, Nahmias et al. (2007) asked participants to attribute free will and moral responsibility to a person who committed a crime. Participants were also exposed either to a 'psychological' scenario (e.g., "*Once specific thoughts, desires, and plans occur in the person's mind, they will definitely*

cause the person to make the specific decision he or she makes”) or to a ‘neuroscientific’ scenario e.g., (“Once specific chemical reactions and neural processes occur in the person’s brain, they will definitely cause the person to make the specific decision he or she makes”) (Nahmias et al., 2007). Results demonstrated that participants reading the neuroscientific scenario attributed significantly less free will and moral responsibility to the wrongdoer as compared to those who read the psychological scenario. While this observation suggests that folk intuitions about free will and moral responsibility can be modulated by different scientific perspectives—in this case a neuroscientific account of human behavior—other researchers have investigated whether weakening people’s general belief about free will has behavioral or cognitive consequences. In other words: If belief in free will is linked to the necessity to hold people responsible for their actions (Baumeister, 2008; Clark et al., 2014), would challenging this belief affect our ethical attitudes?

A number of studies in social psychology have provided consistent answers to this hypothesis. For instance, Vohs and Schooler (2008) found that weakening people’s belief in free will increases cheating. In their first experiment, participants were given a series of mental arithmetic problems that were to be solved. They were told that due to a computer glitch the solution would appear on the computer screen while they were trying to solve the problem, but they were asked to prevent the solution from being displayed by pressing a key after the problem appeared. They were also told that they should solve the problem by themselves, because the experimenter would not know whether they pressed the key. Indeed, the actual dependent variable was the number of times participants pressed the key to stop the answer from appearing, which was taken as a measure of how honest participants’ behavior was. Crucially, before performing the task, participants had read an extract from Francis Crick’s book *The Astonishing Hypothesis* (Crick 1994) that either provided several scientific arguments against free will (e.g., “Although we appear to have free will, in fact, our choices have already been predetermined for us and we cannot change that”) or was neutral about free will. Participants in the ‘anti-free will’ group reported weaker beliefs in free will after reading the Crick text, as measured with a self-report measure of the belief in free will (Paulhus & Carey, 2011). Crucially, they also cheated more frequently than those who were exposed to a neutral message.¹ These data were essentially replicated in a second experiment, in which participants reading the anti-free will text behaved more immorally than the controls in a task involving more active cheating (i.e., stealing money from the researcher; Vohs & Schooler, 2008).

A number of follow-up studies used similar belief manipulations and showed that challenging free will led to a range of antisocial attitudes, including increased aggression toward others, racism, and reduced helpfulness and cooperative attitudes (Baumeister et al., 2009; Protzko, Ouimette, & Schooler, 2016; Zhao, Liu, Zhang,

¹The key finding of Vohs and Schooler’s Experiment 1 was the object of a replication attempt in the Open Science Collaboration project (Open Science Collaboration, 2015). While in the same direction as the original result, the replication result was smaller in effect size and failed to achieve statistical significance.

Shi, & Huang, 2014). Other studies reported a reduction of moral blame and punitive attitudes toward wrongdoers when the belief in free will is weakened (Shariff et al., 2014), suggesting that people also attribute less moral responsibility to others when they have to judge their wrongful actions.

In sum, the studies show that shaking people's belief in free will by means of exposure to a scientific deterministic worldview encourages selfish and impulsive behavior, reduces altruistic attitudes, and increases their tolerance in the face of unethical behavior. Yet, the basic cognitive mechanisms underlying these belief-related changes in social behavior are still poorly understood. It has been proposed that the belief in free will promotes effortful and controlled behavior (Baumeister, 2008); it can therefore be hypothesized that dismissing free will would impact on the cognitive and neural mechanisms on which effortful and controlled behavior is grounded. In the next section, we outline a series of studies reporting how reducing people's belief in free will can influence such basic processes.

Cognitive and Neural Consequences of Disbelieving in Free Will

In the last decades, research in experimental psychology and cognitive neurosciences has started to delineate the cognitive and the neural bases of intentional action control. Theoretical frameworks and models of intentional actions stress the multifaceted nature of voluntary movements (Brass & Haggard, 2008; Brass, Lynn, Demanet, & Rigoni, 2013; Haggard, 2008). For instance, Brass and Haggard (2008) propose that intentional actions involve at least three decisional components concerning the selection of the appropriate action (*what*), the decision about when to perform the action (*when*), and the decision about whether or not the action should be executed (*whether*). Empirical evidence indicates that each of these components relies on distinct brain circuits that involve the pre-supplementary motor area (pre-SMA) and the supplementary motor area (SMA), the rostral cingulate zone (RCz), and the dorsal fronto-median cortex (dFMC) (Brass & Haggard, 2008; Brass et al., 2013). One way to test whether challenging people's belief in free will affects volitional action control is therefore to study how cognitive and neural markers of intentional actions are affected by changes in high-level beliefs about free will. In the following, we outline a series of studies that investigated how reducing people's belief in free will impacts on basic cognitive and neural mechanisms assisting the execution of goal-directed actions.

Increased activity in the pre-SMA and SMA regions is found during tasks involving decisions about when a specific action should be executed (Lau, Rogers, Haggard, & Passingham, 2004; Rigoni, Brass, Roger, Vidal, & Sartori, 2013; Shibasaki & Hallett, 2006). A familiar neurophysiological index of pre-SMA/SMA activity is the readiness potential (RP), a slow electrical potential detectable through electroencephalography (EEG) when the individual performs voluntary hand or

finger movements, such as pressing a key repeatedly at a chosen time (Kornhuber & Deecke, 1965; Shibasaki & Hallett, 2006). The RP is assumed to reflect the activity of the SMA and pre-SMA prior to a voluntary movement (Haggard, 2008; Rigoni et al., 2013; Shibasaki & Hallett, 2006; see Schurger, Sitt, & Dehaene, 2012 for an alternative account). Crucially, this component is sensitive to nonmotor variables such as the level of intention that accompanies the movement (i.e., it is reduced or absent, e.g., for automatic and involuntary movements) and it can therefore be used as an index of the intentional involvement during the execution of an action (Shibasaki & Hallett, 2006).

In an EEG study, Rigoni, Kühn, Sartori, and Brass (2011) investigated whether the RP was sensitive to a reduced belief in free will. The belief manipulation procedure developed by Vohs and Schooler (2008) was employed to weaken people's belief in free will: the anti-free will group reads a scientific text claiming that free will is an illusion, while the control group reads a text on consciousness that was neutral regarding free will. After reading the text, both groups performed a motor task. They were asked to sit in front of a computer screen and to repeatedly press a key at the time of their own choosing. The rationale of this paradigm is that, in the absence of external cues signaling when the action has to be performed, participants have to choose at what time they want to execute the movement (Libet, Gleason, Wright, & Pearl, 1983), resulting in activation of the pre-SMA/SMA brain areas. Self-report measures of individuals' belief in free will indicated that reading the anti-free will text weakened people's belief in free will. Most importantly, the amplitude of the RP preceding the execution of the movement was significantly reduced in the group reading the anti-free will text, as compared to the control group (Rigoni et al., 2011). This observation indicates that dismissing free will attenuates the activation of brain areas that are involved in the decision about *when* to perform a specific action, and was interpreted as evidence that weakening people's belief in free will can affect intentional engagement into action preparation. Further studies expanded the observation that challenging free will affects intentional engagement in motor performance by looking at the subjective experience accompanying intentional actions. Lynn and colleagues (2014) used the intentional binding paradigm to measure individuals' sense of agency over self-produced actions after the belief manipulation (Lynn, Muhle-Karbe, Aarts, & Brass, 2014). They observed that intentional binding was reduced in participants reporting a reduced belief in free will after being primed with an anti-free will scientific text, suggesting that free will beliefs affect people's feeling of being in control of their own actions.

While these studies indicate that weakening abstract beliefs about free will can influence the neural and cognitive indices of intentional action, other studies used similar procedures to investigate whether disbelieving in free will affects components of intentional action more directly related to action control. One crucial aspect of action control is the evaluation of action-effects (Haggard, 2008): efficient action control requires the ability to monitor the consequences of the action and to adjust future actions accordingly. This capacity is particularly important in the case of performance errors: when you write an e-mail, for instance, it is important that you

are able to identify typing errors and correct them. Error monitoring is thus a key component of a broader set of abilities that are typically defined as cognitive control (Botvinick, Braver, Barch, Carter, & Cohen, 2001).

Previous studies have shown that error monitoring processes are malleable to control-related beliefs (Inzlicht & Tullett, 2010; Moser, Schroder, Heeter, Moran, & Lee, 2011). For instance, individuals believing that their abilities and skills can develop through effort (i.e., a belief referred to as growth mindset) show enhanced behavioral and neurophysiological correlates of error monitoring, as compared to individuals who believe they cannot do anything to improve their skills (Moser et al., 2011). In addition, religious believers—i.e., people who believe life events are ultimately controlled by a divine entity—show a reduced neurophysiological response to errors, as compared to nonbelievers (Inzlicht & Tullett, 2010). Taken together, these data show that neurophysiological mechanisms underlying error monitoring can be influenced by individuals' belief about how much they are in control of their own actions and behavior.

One EEG study investigated whether leading people to be skeptical about free will impairs neurophysiological markers of error monitoring and action adjustment (Rigoni, Pourtois, & Brass, 2015). After reading either an anti-free will text or a neutral text, participants performed a go-no-go task, a type of conflict task (Vocat, Pourtois, & Vuilleumier, 2008) in which people are required to respond with a key press as fast as possible to 'go' stimuli (e.g., a green arrow) but must withhold their response to infrequent 'no-go' stimuli (e.g., a black arrow). Time pressure in this type of task typically leads participants to commit a considerable amount of response errors, allowing researchers to investigate the neural signals of error processing. An EEG marker of error processing is the Error-Related Negativity (ERN; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993), a brain wave which is presumably generated by neurons in the midline of the brain, such as the anterior cingulate cortex (ACC) and the SMA (Bonini et al., 2014). The ERN is thought to reflect the comparison between the actual and intended or expected action effect, and it therefore signals the reaction of the brain when responses do not lead to the expected or desired effects (Holroyd & Coles, 2002; Oliveira, McDonald, & Goodman, 2007). In addition, the ERN is sensitive to the subjective value of errors. For instance, the ERN is larger when response errors result in money loss, as compared to when no money is involved (Pailing & Segalowitz, 2004). This component can thus be considered a 'cortical alarm bell' (Inzlicht & Tullett, 2010) that indicates that performance is deviating from desired states. When measured after the anti-free will manipulation, the ERN was significantly reduced as compared to that measured from participants who read a neutral text unrelated to free will (Rigoni et al., 2015). In this study, the ERN was also measured prior to the reading of the anti-free will or neutral texts, and it could be demonstrated that the magnitude of the ERN was significantly diminished in the anti-free will group only, while it remained unaltered in the control group.

While these observations indicate that high-level beliefs about free will can shape neural responses to performance errors, an important question is whether

these effects lead to observable changes in behavioral performance. This hypothesis was tested in a behavioral experiment in which participants performed a similar conflict task before and after reading either anti-free will or neutral scientific material (Rigoni, Wilquin, Brass, & Burle, 2013). Behavioral reactions to errors can be appraised by measuring post-error adaptation (Rabbitt & Rodgers, 1977), that is, a slowing down of the response time on a given trial n after committing an error on the previous trial $n-1$. Post-error adaptation is assumed to reflect the implementation of control mechanisms intended to adjust subsequent behavior after an error, and therefore signals increased ‘response caution’ following an error (Dutilh et al., 2012; see Notebaert et al., 2009 for an alternative account). While large post-error adaptation indicates the implementation of cognitive control, a diminished post-error adaptation reveals degraded action monitoring (Botvinick et al., 2001). It was observed that post-error adaptation was diminished in participants reading the anti-free will text, as compared to control participants. Leading people skeptical about free will diminished response caution after an error was committed, presumably indicating a careless attitude toward performance improvement (Dutilh et al., 2012).

In sum, these studies provide evidence that basic cognitive and neural markers of action control are altered when participants’ belief in free will has been challenged. These observations substantiate the interpretation that believing in free will encourages the implementation of self-regulatory mechanisms, such that reducing free will beliefs can decrease the willingness to implement behavioral control.

Conclusion and Final Remarks: Is Belief in Free Will a Meta-Representation?

Here, we presented empirical evidence indicating that philosophical intuitions concerning free will are grounded in the bodily experience of being in control of ourselves (Aarts & van den Bos, 2011; Ent & Baumeister, 2014), and that the extent to which people believe in free will depends, in part, on the desire to hold people responsible for their own behavior (Clark et al., 2014; Knobe & Fraser, 2008; Young & Phillips, 2011). At the same time, scientific arguments that challenge the existence of free will can alter people’s belief in free will, which in turn leads to changes in social behavior (Baumeister et al., 2009; Shariff et al., 2014; Vohs & Schooler, 2008) as well as cognitive (Lynn et al., 2014; Rigoni, Kühn, Gaudino, Sartori, & Brass, 2012; Rigoni et al., 2013) and neural (Rigoni et al., 2011, 2015) markers of action control.

These data suggest that discussions around the basis of actions can alter our experience and can change our behavior. It remains unclear, however, how abstract beliefs about free will and related constructs influence (and are influenced by) social behavior as well as neural and cognitive mechanisms of action control. We speculate that the belief in free will can be conceptualized as a form of metacognitive

judgment of actions of the self and others (Frith, 2012). While metacognition is typically defined in a very broad sense as the ability to consciously reflect upon one's own states, or 'cognition about cognition' (Carruthers, 2009), more formalized models define it more narrowly as an unconscious representational *re-description* process (Timmermans, Schilbach, Pasquali, & Cleeremans, 2012) or as a set of control processes that make use of representations of the properties of other cognitive processes (Shea et al., 2014). The belief in free will can therefore be viewed as a meta-representation that defines how lower level or first-order representations are implicitly or explicitly held and processed (Timmermans et al., 2012). In addition, as any other belief, the representational content of the belief in free will can be described as a propositional attitude about how we (and others) act in the environment (e.g., "I can decide to do X even in presence of internal or external constraints") and how we perceive and interact with others (e.g., "He deserves to be punished because he acted freely") (Connors & Halligan, 2014).

It can be assumed that, if the belief in free will influences how we act, it would also affect how we perceive and judge others' behavior (Gallese, 2007). The concept of free will thus links one's own subjective experience of actions that we cause and one's perception of other people's actions. In this respect, it is interesting to note that contemporary models of metacognition (Fleming, Dolan, & Frith, 2012) explicitly assume that the mechanisms through which one assesses one's own performance are germane with the mechanisms involved in appraising other people's behavior. Metacognition, in this sense, is intimately linked to theory of mind (i.e., the ability to attribute mental states, such as intentions, emotions and beliefs to others). In both cases, high-level redescrptions of first-order behavior provide a narrative through which to understand the basis of action. Our proposal is therefore that such high-level redescrptions are subtended by internal models of what it takes to be a free agent, and as such, they can be readily modified by other beliefs, by social interactions, and by the specific circumstances in which action is deployed or observed. There is thus an interactive, embodied loop that extends not only from our behavior to our subjective experience of action, but also from the behavior of other people to our judgments about that behavior.

One can therefore speculate that explicit beliefs about free will would affect social cognition, namely how individuals conceptualize their social environment. Previous findings indicate that weakening people's belief in free will reduces punishing attitudes toward hypothetical criminals (Krueger, Hoffman, Walter, & Grafman, 2014; Shariff et al., 2014). These results show that disbelieving in free will influences how people interpret others' behavior. One possible interpretation of these data is that a decreased belief in free will redescrbes the representation of our own and others' behavior as more constrained by internal or external contingencies. Further research should thus examine how more basic cognitive and neurophysiological processes of social cognition are influenced by people's (dis)belief in free will.

References

- Aarts, H., & van den Bos, K. (2011). On the foundations of beliefs in free will: Intentional binding and unconscious priming in self-agency. *Psychological Science*, 22(4), 532–537.
- Baumeister, R. F. (2008). Free will in scientific psychology. *Perspectives on Psychological Science*, 3, 14–19.
- Baumeister, R. F., Masicampo, E. J., & Deway, C. N. (2009). Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin*, 35, 260–268.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624–652.
- Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist*, 14(4), 319–325.
- Brass, M., Lynn, M. T., Demanet, J., & Rigoni, D. (2013). Imaging volition: What the brain can tell us about the will. *Experimental Brain Research*, 229(3), 301–312.
- Carey, J. M., & Paulhus, D. L. (2013). Worldview implications of believing in free will and/or determinism: Politics, morality, and punitiveness. *Journal of Personality*, 81(2), 130–141.
- Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32(02), 121–138.
- Clark, C. J., Luguri, J. B., Ditto, P. H., Knobe, J., Shariff, A. F., & Baumeister, R. F. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology*, 106, 501–513.
- Connors, M. H., & Halligan, P. W. (2014). A cognitive account of belief: A tentative road map. *Frontiers in Psychology*, 5, 1588.
- Crescioni, A. W., Baumeister, R. F., Ainsworth, S. E., Ent, M., & Lambert, N. M. (2015). Subjective correlates and consequences of belief in free will. *Philosophical Psychology*, 29(1), 41–63.
- Crick, F. (1994). *The astonishing hypothesis: The scientific search for the soul*. New York, NY: Touchstone.
- Dennett, D. C. (2003). *Freedom evolves*. New York, NY: Viking.
- Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E.-J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, 74(2), 454–465.
- Ent, M. R., & Baumeister, R. F. (2014). Embodied free will beliefs: Some effects of physical states on metaphysical opinions. *Consciousness and Cognition*, 27, 147–154.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78(6), 447–455.
- Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1594), 1280–1286.
- Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1599), 2213–2223.
- Gallagher, S. (2012). Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30(1), 15–31.
- Gallese, V. (2007). Before and below ‘theory of mind’: Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480), 659–669.
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, 4(6), 385–390.
- Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, 9(12), 934–946.

- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382–385.
- Harris, S. (2012). *Free will*. New York, NY: Free Press.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679–709.
- Inzlicht, M., & Tullett, A. M. (2010). Reflecting on God: Religious primes can reduce neurophysiological response to errors. *Psychological Science*, 21(8), 1184–1190.
- Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.
- Kashima, Y., Yamaguchi, S., Kim, U., Choi, S. C., Gelfand, M. J., & Yuki, M. (1995). Culture, gender, and self: A perspective from individualism-collectivism research. *Journal of Personality and Social Psychology*, 69(5), 925–937.
- Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (pp. 441–448). Cambridge, MA: MIT Press.
- Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv Für Die Gesamte Physiologie Des Menschen Und Der Tiere*, 284, 1–17.
- Krueger, F., Hoffman, M., Walter, H., & Grafman, J. (2014). An fMRI investigation of the effects of belief in free will on third-party punishment. *Social Cognitive and Affective Neuroscience*, 9(8), 1143–1149.
- Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science*, 303(5661), 1208–1210.
- Laurin, K. (2017). Religion and its cultural evolutionary by-products. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, 106, 623–642.
- Lynn, M. T., Muhle-Karbe, P. S., Aarts, H., & Brass, M. (2014). Priming determinist beliefs diminishes implicit (but not explicit) components of self-agency. *Frontiers in Psychology*, 5, 1483.
- Miller, J. G., & Bersoff, D. M. (1992). Culture and moral judgment: How are conflicts between justice and interpersonal responsibilities resolved? *Journal of Personality and Social Psychology*, 62(4), 541–554.
- Monroe, A. E., & Malle, B. F. (2010). From uncaused will to conscious choice: The need to study, not speculate about people’s folk concept of free will. *Review of Philosophy and Psychology*, 1(2), 211–224.
- Monroe, A. E., Dillon, K. D., & Malle, B. F. (2014). Bringing free will down to Earth: People’s psychological concept of free will and its role in moral judgment. *Consciousness and Cognition*, 27, 100–108.
- Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, 21(1), 546–561.
- Moser, J. S., Schroder, H. S., Heeter, C., Moran, T. P., & Lee, Y.-H. (2011). Mind your errors: Evidence for a neural mechanism linking growth mindset to adaptive post-error adjustments. *Psychological Science*, 22(12), 1484–1489.
- Nahmias, E., Coates, D. J., & Kvaran, T. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy*, 31(1), 214–242.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology*, 18(5), 561–584.
- Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, 111(2), 275–279.

- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: Expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, *19*(12), 1994–2004.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).
- Pailing, P. E., & Segalowitz, S. J. (2004). The error-related negativity as a state and trait measure: Motivation, personality, and ERPs in response to errors. *Psychophysiology*, *41*(1), 84–95.
- Paulhus, D. L., & Carey, J. M. (2011). The FAD-Plus: Measuring lay beliefs regarding free will and related constructs. *Journal of Personality Assessment*, *93*(1), 96–104.
- Pronin, E., & Kugler, M. B. (2010). People believe they have more free will than others. *Proceedings of the National Academy of Sciences*, *107*(52), 22469–22474.
- Protzko, J., Ouimette, B., & Schooler, J. (2016). Believing there is no free will corrupts intuitive cooperation. *Cognition*, *151*, 6–9.
- Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *The Quarterly Journal of Experimental Psychology*, *29*(4), 727–743.
- Rigoni, D., Brass, M., Roger, C., Vidal, F., & Sartori, G. (2013a). Top-down modulation of brain activity underlying intentional action and its relationship with awareness of intention: An ERP/Laplacian analysis. *Experimental Brain Research*, *229*(3), 347–357.
- Rigoni, D., Kühn, S., Gaudino, G., Sartori, G., & Brass, M. (2012). Reducing self-control by weakening belief in free will. *Consciousness and Cognition*, *21*(3), 1482–1490.
- Rigoni, D., Kühn, S., Sartori, G., & Brass, M. (2011). Inducing disbelief in free will alters brain correlates of preconscious motor preparation: The brain minds whether we believe in free will or not. *Psychological Science*, *22*, 613–618.
- Rigoni, D., Pourtois, G., & Brass, M. (2015). “Why should I care?” Challenging free will attenuates neural reaction to errors. *Social Cognitive and Affective Neuroscience*, *10*(2), 262–268.
- Rigoni, D., Sammiceli, L., & Sartori, G. (2015). Looking for the right intention: Can neuroscience benefit from the law. *Frontiers in Human Neuroscience*, *9*(432).
- Rigoni, D., Wilquin, H., Brass, M., & Burle, B. (2013b). When errors do not matter: Weakening belief in intentional control impairs cognitive reaction to errors. *Cognition*, *127*(2), 264–269.
- Sarkissian, H., Chatterjee, A., De Brigard, F., Knobe, J., Nichols, S., & Sirker, S. (2010). Is belief in free will a cultural universal? *Mind and Language*, *25*(3), 346–358.
- Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(42), E2904–E2913.
- Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J. B., Clark, C. J., Schooler, J. W., et al. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science*, *25*(8), 1563–1570.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Sciences*, *18*(4), 186–193.
- Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology*, *117*(11), 2341–2356.
- Timmermans, B., Schilbach, L., Pasquali, A., & Cleeremans, A. (2012). Higher order thoughts in action: Consciousness as an unconscious re-description process. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1594), 1412–1423.
- Vocat, R., Pourtois, G., & Vuilleumier, P. (2008). Unavoidable errors: A spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia*, *46*(10), 2545–2555.
- Vohs, K. D., & Schooler, W. (2008). The value of believing in free will. *Psychological Science*, *19*, 49–54.
- Young, L., & Phillips, J. (2011). The paradox of moral focus. *Cognition*, *119*(2), 166–178.
- Zhao, X., Liu, L., Zhang, X., Shi, J., & Huang, Z. (2014). The effect of belief in free will on prejudice. *PLoS ONE*, *9*, e91572.

Religion and Its Cultural Evolutionary By-products

Kristin Laurin

Lay theories are the beliefs and theories that people have about everyday life—be that their own selves, others’ behavior, or events in the world. Any volume on lay theories would therefore be incomplete with a chapter on one of the most culturally prevalent set of beliefs about life in all of human history: religion. Religious beliefs in some form or another have been present in human societies from as early as 100,000 years ago (Greenspan, 2006; Lieberman, 1991), and continue to play a role in conflicts—both violent and political—in today’s modern world. In terms of individual psychology, 84% of Americans—and similarly, 84% of people worldwide—identify with a particular religion (Pew Research Center, 2012). Recognizing the fundamental nature of religion, the recent psychological literature has begun to delineate the broad swatch of human life touched by religion and religious beliefs. Here, I focus on one type of religious belief in particular—the belief in God—and more specifically on its consequences for individual human functioning. In the first part of this chapter, I describe one prominent theory of why we tend to believe in the kinds of Gods we believe in, one that relies on principles of evolution applied to the level of cultures. In the second part, I outline some of the cultural evolutionary by-products of these beliefs, and discuss the implications for religion’s continued evolution.

K. Laurin (✉)

Department of Psychology, University of British Columbia,
2136 West Mall, Vancouver, BC V6T 1Z4, Canada
e-mail: klaurin@psych.ubc.ca

PART I: Identifying and Explaining God Beliefs

What Is Religion and What Is God?

Scholars have identified four major dimensions that can be used to describe any modern religion (Saroglou, 2011). First, religions have a *behavioral* component: Part of what it means to be religious is to follow a set of prescribed rules. Some religions impose dietary restrictions on their adherents, for instance, and others demand that they follow a strict schedule of daily prayer. Second, religions typically involve *communities*: Most often, at least some of the behaviors are practiced in the company of others. Members of these communities can monitor each other's compliance with the religion's rules, and reinforce each other's commitment to the faith. Third, religions provide their adherents with a *moral code*: a set of guidelines that define what is right and what is wrong. Fourth, religions organize and legitimize all three of these components through a system of shared *beliefs*. In particular, central to many modern religions—and central to the aims of this chapter—is the belief in a powerful supernatural force often called God.

But what, exactly, is God? Our species has seen hundreds of different religions, and they have each treated the idea of God differently. Polytheistic religions worship numerous Gods—for example, Hindus pray to a large number of deities—while monotheistic religions, like Judaism, Islam and Christianity, believe in a single supreme being. Different monotheistic religions represent this supreme being somewhat differently—for instance Christians pray to a singular God that is supposed to exist in three different forms (Father, Son and Holy Spirit), whereas Jews and Muslims believe in a God who has only one form. And some religions, like Buddhism, appear to not believe in any sort of God at all.

Nevertheless, from these differences there emerge some commonalities, and in particular, many scholars argue that most religions in practice draw people toward believing in some form of singular God-like entity. For instance, in polytheistic religions like Hinduism, many followers adopt a view more akin to *polymorphism*: They may worship Gods that take on different forms, while also believing that all those forms emanate from a single divine essence. And in seemingly nontheistic religions like Buddhism, adherents often in practice believe in a singular creating God-like essence (see Norenzayan, 2013; Norenzayan & Atran, 2004, for a review).

Thus, the religions that are most widely practiced today gravitate toward a belief in a single God; moreover, they tend to attribute that God the same three characteristics. The Gods most people worship today are powerful—that is, they can intervene in worldly affairs and directly influence human outcomes. They are watchful—that is, they can observe worldly affairs and are very aware of human actions. And they are morally invested—that is, they care about how humans behave, and prefer them to behave morally, rather than immorally. These characteristics are most clearly true of the world's major monotheistic religions (Christianity, Judaism and Islam); however, the notion of a powerful supernatural

being capable of punishing humans for misbehaving exists even in societies and religions traditionally thought to be atheistic (e.g., in China, Slingerland, 2013; and in ancient Greece and Rome; Mikalson, 2010; Rives, 2007).

The Cultural Evolution of Big God Beliefs

Several different perspectives exist to explain why human beings tend to believe in God. Some have argued that beliefs in God are a by-product of the way humans think (Atran & Norenzayan, 2004; Boyer, 2001; Barrett, 2000). For instance, we may believe in God because of our tendency to mentalize (e.g., Barrett & Keil, 1996) or to believe that events happen and objects exist for a purpose (e.g., Evans, 2001; Pyysiäinen, 2009). Others suggest that we believe in God because it assuages our existential anxieties, or fulfills other important intrapsychic needs (e.g., Gray & Wegner, 2010; Vail, Arndt, & Abdollahi, 2012). The perspective I take in this chapter, though, is based on the notion of *cultural evolution*.

The Cultural Evolution Account

According to this perspective, it is no mere coincidence that so many different human societies embrace not only beliefs in a supernatural agent, but seemingly extremely similar beliefs about supernatural agents who are specifically powerful, watchful, and morally invested. Indeed, some recent scholarship suggests that beliefs in these kinds of agents emerged through a process of cultural evolution, spreading because they solve a pressing human need: The need to maintain cooperation in large-scale groups. Genetic (Boyd & Richerson, 1988; de Waal, 2008) and reputational mechanisms (e.g., Henrich & Henrich, 2007; Panchanathan & Boyd, 2003) can explain why we cooperate with our family members, and with others with whom we have ongoing relationships. But once human societies expand beyond small groups where everyone knows everyone else—that is, beyond groups of about 150 people (Dunbar, 2003; but see Smith, 1996)—those mechanisms can no longer account for how we manage to, for the most part, get along and cooperate with each other (Chudek & Henrich 2011; Chudek, Brosseau, Birch, & Henrich, 2013).

Instead, the cultural evolution hypothesis proposes that, to explain large-scale human societies, we need mechanisms through which people believe that their behavior will be monitored, and potentially punished, by a force more powerful than other humans. Powerful, watchful, morally invested Gods—otherwise known as Big Gods (Norenzayan, 2013)—provide just such a mechanism. If I believe in a Big God, then I know that if I cheat my neighbor, even if I do it under the cover of anonymity, the Big God will see it (because he is watchful), judge me negatively (because he is morally concerned) and punish me (because he is powerful enough to do so). Therefore, if everyone in my group believes in such a God, then everyone

will cooperate. This, according to the cultural evolution hypothesis, is the cultural evolutionary pressure that selected for Big God beliefs: Cultures that promoted shared beliefs about powerful, watchful, morally invested Gods survived and thrived because they could grow while continuing to cooperate; those that did not could not grow beyond a certain size without dissolving in conflict or other failures of cooperation.

Some Evidence for the Cultural Evolution Account

As evidence for this theory, scholars point to different sorts of evidence. One set of empirical findings consists of anthropological observations relating various features of groups to the presence of Big Gods. For instance, even controlling for numerous potential confounds, cross-cultural analyses find that the size of a society directly predicts the “size” of that society’s Gods (Peoples & Marlowe, 2012; Roes & Raymond, 2003). That is, larger societies are more likely to contain cultural notions of a single powerful, watchful, and morally concerned God. As another example, cultures where natural resources are scarce are more likely to have Big Gods who are believed to specifically encourage people to be cooperative and prosocial in their use of these resources (Botero et al., 2014). This is consistent with the argument that Big God beliefs are an important cultural tool for solving the problem of large-scale cooperation, and therefore for permitting societies both to expand and also to subsist in suboptimal environments.

A Digression About Priming Religious Beliefs

Another critical set of empirical findings is that Big Gods in today’s modern society have three effects consistent with the idea that they continue to serve the culturally adaptive function of promoting broad cooperation. These effects—along with many of the findings discussed in this chapter—have largely been documented using a technique called *priming*, so it is worth a brief digression to describe how this works. *Priming* involves activating a mental representation in someone’s mind. If I talk to you about chocolate cheesecake, for instance, this will activate your mental representation of chocolate cheesecake, because that is what you will be thinking about. But it may also activate other related mental representations—for instance your mental representations of chocolate bars, of strawberry shortcake, of dessert in general, and maybe even of that diet you promised yourself you would stick to (Schvaneveldt & Meyer, 1973). But the key to priming is that once activated, a mental representation remains accessible for some duration, and can thereby influence a person’s perceptions, judgments, and behavior (e.g., Eitam & Higgins, 2010). To use the same example, for some period of time after our conversation, if you see something brown and round you may mistake it for a chocolate cake, or you may find yourself reaching for a chocolate bar when you get hungry, rather than a different snack option. In the context of research on Big Gods, researchers

have primed the concept of God, sometimes explicitly (e.g., by having participants read a passage about God, e.g., Laurin, Kay, & Fitzsimons, 2012a, Study 6), or by having them answer questions about their beliefs about God (e.g., Laurin, Shariff, Henrich, & Kay, 2012b), and sometimes implicitly (e.g., by having participants play a word game that happens to involve a few words related to God in it—one advantage of implicit priming techniques is that they preempt demand characteristics: Because participants are not aware that the researchers intended to make them think about God, they cannot guess how the researchers hope they will act, and this guessing can therefore not influence their behavior).

By comparing the behavior of a group of people who have completed this kind of priming task to the behavior of a different group of people assigned to a control task, researchers can infer the causal influence of having God on the brain. That is, they can identify how religious thoughts influence people's thoughts or behavior. The real question of this chapter, though, is about the role of religious *beliefs*, not religious thoughts. The (methodological) problem with religious beliefs is that they are not randomly distributed, so any kind of thought or behavior that is more prevalent in believers than nonbelievers is not necessarily *caused* by belief. The most illuminating studies on this front have therefore combined priming techniques with measuring people's own beliefs to uncover the beliefs' causal role. For instance, some use a priming manipulation that simply varies whether or not participants are asked about their beliefs about God prior to completing the experiment, and observe effects that emerge only among participants who are thus reminded of the fact that *they themselves* believe strongly in Big Gods (e.g., Laurin et al., 2012b). Studies that use other priming methods tell us most directly about the causal role of religious thoughts. However, it is extremely likely that religious beliefs are a source of frequent religious thoughts: If I believe in God, I will likely find myself thinking about God more often than someone who does not. Moreover, I will probably surround myself with people who share my belief (e.g., Hitsch, Hortacsdu, & Ariely, 2010), and the conversations I have with these people may cause me further to reflect on God. Therefore, the consequences of thinking about God are very likely, also consequences of believing in God, although beliefs may have additional consequences that priming methods are less suited to discovering.

More Evidence for the Cultural Evolution Account

To return to the subject at hand, the priming techniques I just described are what have allowed researchers to demonstrate three effects of Big God beliefs that support the idea that these beliefs continue to serve the culturally adaptive function of promoting broad cooperation. First, people seem to equate Big Gods with monitoring: Thinking about these Gods makes people feel they are being watched (Gervais & Norenzayan, 2012). Importantly, they seem to believe Big Gods are specifically watching for and noting violations of norms or group morals: People are quicker to respond when asked whether Big Gods know about norm-violating behaviors, than when asked whether they know about other kinds of behaviors,

indicating that they more readily associate Big Gods with knowing about norm violations (Purzycki et al., 2012). Thus, people's current interactions with Big Gods are consistent with the notion that these Gods emerged in part because they make people feel that their behaviors—in particular their behaviors that go against group norms, like failing to cooperate—will be seen.

Second, presumably in part thanks to this monitoring function, Big Gods do help ensure cooperation. It has now been nearly ten years since two independent labs working separately produced the first experimental evidence demonstrating that religious notions, and more specifically the idea of God, can make people less likely to cheat and more likely to be generous toward strangers. For example, in one study, participants who first read a series of words related to religion were less likely to cheat by opening their eyes during a money making task they were meant to do with their eyes closed (Randolph-Seng & Nielsen, 2007). In another, participants who played a word game that involved words related to God—i.e., primed with the notion of God—shared more of the money the experimenter gave them with an anonymous other participant (Shariff & Norenzayan, 2007). In other words, the notion of God can make people less likely to selfishly prioritize their own gain at the expense of others (e.g., of the experimenter who had to pay the cheaters extra money), and more likely to do things that will instead help others. Follow-up work has found that these effects emerge primarily because of people's views of God as punitive—in other words, God does not make people nice because they believe in a nice God and they want to be like God; rather, God makes people nice because they want to avoid God's punishment (Shariff & Norenzayan, 2011; Yilmaz & Bahçekapili, 2016; see also Henrich et al., 2010). A recent review of the nearly 100 studies that have subsequently replicated this basic pattern indicates that these effects are reliable and replicable (Shariff, Willard, Andersen, & Norenzayan, 2016). Together, then, substantial empirical evidence confirms that Big Gods can ensure cooperation.

Third, and finally, Big Gods help attenuate favoritism toward close others, and broaden the scope of people's generosity. For instance, in one set of studies conducted in the United States, priming participants with the notion of God, but *not* priming them with the concept of religion in general, made them more likely to donate money to the Mexican Red Cross—that is, to share their resources with a group dedicated to helping outgroup members (Preston & Ritter, 2013). Similar results have also been found using villagers from a small-scale horticulturalist society: Villagers who more strongly believed in punitive Gods were more likely to give money to a random stranger from a different village, rather than to themselves or to a resident of their own village (McNamara, Norenzayan, & Henrich, 2016). In other words, Big Gods may help people to be less parochial, and to instead incorporate strangers into their group, which is consistent with the idea that Big Gods played a critical role in helping societies expand.

PART II: By-products of Culturally Evolved Big Gods Beliefs

Thus, evidence is mounting for the notion that beliefs in Big Gods emerged as a cultural adaptation to solve a specific societal need. However, this does not mean that the effects of beliefs in Big Gods are constrained to only relate to that specific societal need. Features that evolve because they serve one particular adaptive function can nevertheless serve other functions, or have other effects, as well.

Examples of such evolutionary by-products abound in the literature on genetic evolution, and they can be either positive or negative. When the by-products have an adaptive effect, they are sometimes called exaptations or spandrels (Buss, Haselton, Shackelford, Bleske, & Wakefield, 1998). As an example of such an adaptive by-product, take birds' ability to fly (e.g., Sumida & Brochu, 2000). The latest scholarship indicates that feathers, which are critical to birds' ability to fly, did not in fact evolve because of that function. Instead, scientists believe that feathers may have evolved because they help animals regulate their temperatures, protect themselves from water, or dispose of waste (Bock, 2000). Only later did some organisms start using their feathers for flight, and feathers have since evolved further in response to this new function, such that they now provide more flight-related functionality than temperature-related functionality. Importantly, this example illustrates how it may be possible for evolved features to stop serving their original function, if they begin to serve a new, still adaptive function: It seems at least possible that feathers could lose their heat regulatory function altogether, and still continue to exist because they enable flight, which is crucial to most birds' survival.

Other times, though, the by-products are negative. For instance, consider sickle cell anemia (e.g., Williams et al., 2005). People whose genetic background traces back to geographical regions where malaria has long been rampant often have a particular genetic adaptation that changes the shape of their red blood cells to protect them against the disease. But at the same time, this adaptation—called sickle cell—causes anemia. Evolutionary processes nevertheless retained the sickle cell genes because anemia, the negative side effect, is less severe than malaria, against which the genes provide a defense.

Thus, at least when it comes to genetic evolution, we know that just because a feature evolved as a result of one particular adaptive function, it may continue to persist while causing other effects as well, whether these are adaptive, maladaptive, or neutral. This phenomenon has not been formally studied in the domain of cultural evolution, perhaps because the scholarship on cultural evolution is still in its infancy: Although early thinkers acknowledged that evolution might occur at the level of groups (Darwin, 1877; James, 1880), the first scientists to take this idea seriously and develop it into a more complete theory did not do so until the 1980s (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981; Lumsden & Wilson, 1981). That said, it seems straightforward to assume that cultural evolution can generate by-products in a similar fashion to genetic evolution. Any characteristic,

like beliefs in Big Gods, that has evolved as a result of a particular adaptive need may nevertheless have additional effects, unrelated to that need, that are either positive or negative—so long as its net effect remains adaptive for the broader culture. In the remainder of this chapter, I consider each of the traits of Big Gods, and discuss findings on their by-products.

By-products of Big Gods' WATCHFUL Nature

First, what might be the by-products of the watchful nature of Big Gods? The existing literature points to one particular domain outside of cooperation where watchful Gods may influence human behavior: self-regulation (see Laurin & Kay, 2016). Briefly, self-regulation means guiding and correcting one's behavior in the service of pursuing goals and achieving psychological well-being (Barkley, 1997; Carver & Scheier, 2011; Kanfer & Karoly, 1972; Muraven & Baumeister, 2000). Moreover, self-regulation rests on the convergence of two different factors: Good self-regulation requires a general set of abilities (Carver & Scheier, 1998), but it also requires motivation, or a situation-specific willingness to engage self-regulatory skills (Bandura, 1997; Vroom, 1964).

Watchful Gods and Self-regulatory Ability

In theory, Big Gods should help boost both of these elements, although the empirical evidence is merely suggestive rather than conclusive. First, in terms of the abilities required for self-regulation, one critical skill is self-monitoring: Once a student has settled on her target GPA, she must continually note her performance and compare it to the standard she has set if she is to detect the discrepancies and act to reduce them (Carver & Scheier, 1998). All else being equal, then, she will have better self-regulatory skills if she more regularly checks in and notes her performance.

I noted earlier that Big Gods' watchful nature can enhance public self-awareness: That is, thinking about Big Gods can make people extra conscious of how they appear to others. Moreover, more recent findings indicate that thinking about God can prompt private self-awareness (Kitchens, 2015). Both these states imply some form of monitoring, with public self-awareness making people aware of their external states related to interpersonal goals, and private self-awareness making people aware of their internal states related to intrapersonal goals (Arkin & Baumgardner, 1986; Baumeister, 1982; Schlenker, 1980, 1986). Thus, to return to our example, if the student comes to think about God, she will find herself regularly updating her sense of where she stands—whether relative to her own internal standards or to her perception of others' standards for her. As a result, she will more often note discrepancies that she needs to act on. In other words, because Big Gods are watchful, they may boost self-monitoring, which will in turn make people better

at self-regulation. Some research supports this idea (Carter, McCullough, & Carver, 2012; Kim-Spoon, Farley, Holmes, Longo, & McCullough, 2014; see also McCullough & Willoughby, 2009). However, this research has not examined the causal role of Big Gods' watchfulness specifically, and has rather examined the correlational associations of religiosity more broadly. There may be other reasons why religion promotes self-monitoring, and thus self-regulation; future research may isolate the role of Big Gods.

Watchful Gods and Self-regulatory Will

Second, in terms of the motivation required for self-regulation, Big Gods' watchfulness can help as well. In particular, the sense that one is being watched, or that others are monitoring one's behavior, can instill in people a desire to avoid misbehaving (Baldwin & Holmes, 1987; Dahl, Manchanda, & Argo, 2001; Latané, 1981; Leary, 1995; see also Zhong, Bohns, & Gino, 2010). In the context of self-regulation, "misbehaving" means succumbing to temptation: If I am on a diet, eating the chocolate cake instead of the fruit salad for dessert constitutes misbehaving. Thus, Big Gods' watchfulness, because it makes people feel monitored by others, should motivate them to resist temptations.

The cultural evolution hypothesis and the evidence for it that I described in the first part of this chapter already demonstrates how Big Gods and their watchfulness can help people resist specific temptations to cheat or to be selfish (e.g., Randolph-Seng & Nielson, 2007; Shariff & Norenzayan, 2007)—indeed this may be the primary function that Big Gods emerged to serve. The argument I am making here, though, is that in addition to that culturally adaptive benefit of helping societies avoid the temptation of behavior that is not prosocial and cooperative, these Big Gods provide the by-product of helping people avoid temptations in a broad range of domains.

Some studies support this notion as well. For instance, in one study of participants pursuing health goals, those who read a passage about God ate fewer cookies in the context of a taste testing session, compared to those who read an unrelated neutral passage (Laurin et al. 2012a). Not only that, in a second study in the same paper, participants who played a word game filled with God-related words showed more negative implicit associations with unhealthy foods on a later reaction time task, relative to participants who played a word game that included no God-related words.

To summarize, perhaps because thinking about a watchful God makes people monitor themselves, the notion of God can make people skilled at pursuing their goals. Moreover, at least in part because thinking about a watchful God makes people feel that others are monitoring them, the notion of God also makes them willing to resist the temptations that threaten to derail their pursuits, even when it comes to goals that are unrelated to the kinds of cooperative, prosocial goals that Big Gods presumably evolved to facilitate.

By-Products of Big Gods' POWERFUL Nature

Second, what might be the by-products of the powerful nature of Big Gods? Here the research points to effects in two separate domains.

Self-regulation

One of these domains is, again, self-regulation, but the effects of a powerful God are less straightforward than the effects of a watchful God (see Laurin & Kay, 2016). On the one hand, a powerful God may decrease people's motivation—if not their skill—to pursue their goals. One of the critical beliefs that is required for motivation to exist in the first place is the belief that one can, with sufficient effort, accomplish the goal (e.g., Bandura, 1997). But, as might be expected given God's powerful nature, when believers are exposed to the word "God"—even very briefly, so briefly that they could not have consciously noticed it—they feel they play less of a role in causing events that occur (Dijksterhuis, Preston, Wegner, & Aarts, 2008). That is, after being subliminally primed with the word "God", people claimed that they were less responsible for making a letter disappear from a computer screen, presumably because they attributed some of that responsibility to God. Similarly, a theoretical model of compensatory control (Kay, Gaucher, Napier, Callan, & Laurin, 2008) suggests that perceptions of God's control have a hydraulic relationship with perceptions of personal control: The more control a person ascribes to herself, the less she ascribes to God, and vice versa. Together, these ideas suggest that God's powerful nature may make people feel that they themselves lack the ability to produce the outcomes they desire, which can reduce their motivation to try to pursue these outcomes.

A similar argument can be made drawing on the social loafing literature (Karau & Williams, 1993). Research on social loafing shows that when people share the responsibility for their outcomes with other entities, they recognize that their outcomes are therefore contingent on not only their own but also others' decisions and behaviors. As a consequence, sharing the responsibility in this way makes people invest less effort toward goal attainment (Karau & Williams, 1993). To use a classic example, imagine two different tug of wars: One pitting two individual people against each other, and one where each of those people has five teammates also pulling with them. The social loafing literature suggests that often, people will pull less hard in the second scenario, relying on their teammates to do most of the pulling. Similarly, if people view a powerful God as being on their team—as helping them to achieve their goals—they may lessen their own efforts, relying on God to help them succeed.

Thus, if people view a powerful God as constraining their own ability to control their outcomes, or as intervening on their behalf to help them succeed, then this powerful God may decrease people's willingness to invest their own efforts in pursuit of important goals. Our own research has documented this very

phenomenon (Laurin et al., 2012a). For instance, engineering students who played a word game that involved a few words related to God later worked less hard on an anagram task they believed was related to their likelihood of success in engineering. Moreover, they were *especially* likely to work less hard if they had previously told experimenters that they believe that external forces could possibly influence their engineering success: There was no effect at all among participants who had instead indicated that they felt they were the sole authors of their professional destiny. And, in a different study in the same paper participants who read a passage about a powerful God—but not those who read about a watchful God or a creator God—expressed less of a willingness to invest efforts in pursuing their self-identified career goals. Both these findings suggest that the thought of a powerful God who can influence one's outcomes makes people work less hard in pursuit of that particular outcome.

A different theoretical perspective, though, suggests that the idea of a powerful, interventionist God can instead *increase* people's motivation to pursue their goals. External controlling forces like Big Gods provide order and structure in our daily life, ensuring that the world is a sensible place where specific actions do in fact lead to predictable outcomes. From this perspective, a powerful God can guarantee the kind of order and contingency that is a prerequisite for motivation (Kay, Laurin, Fitzsimons, & Landau, 2014; Landau, Kay, & Whitson, 2015): In order to be willing to try hard and invest efforts in pursuit of goals, people must have faith that specific consequences reliably follow certain actions (Vroom, 1964; Laurin, Fitzsimons, & Kay, 2011).

Some evidence supports this perspective as well: In five experimental studies, people invested more effort in pursuit of their goals after thinking about environmental order—even order unrelated to the goals they were pursuing (Kay et al., 2014). Similarly, disorganized environments hindered self-regulation, whereas organized environments resulted in better self-regulation (Chae & Zhu, 2014). Although none of this research directly examined the role of powerful Gods, if it is true that these Gods can help restore people's sense that the world is orderly (Kay et al., 2008), they suggest that these Gods can thereby facilitate goal pursuit.

Thus, two separate lines of work indicate that powerful Gods can have opposing effects on people's motivation to pursue their goals. On the one hand, they can make people feel they have less control over their outcomes, or that a powerful friend is stacking the deck in their favor, which can make them try less hard. On the other hand, powerful Gods can make people feel confident that the world is orderly and sensible, which is required for them to be willing to invest their own time and energy, and which may make them try harder. No published papers have reconciled between these two sets of findings, but emerging evidence suggests a possible moderator that can determine whether people see God as a welcome provider of order and structure, or instead as a force capable of interfering with or substituting for their own power. In particular, these perceptions may depend on people's current sense of their own ability to pursue their goals. If they feel unsure in this regard, a powerful God may remind them that the world is an orderly, logical place, and that all they have to do is take things one step at a time. If instead they are

feeling confident in their own individual abilities, then a powerful God may remind them that their own individual abilities are not all that matters. Consistent with these ideas, when people are made to question their ability to succeed—for instance through manipulations that reduce self-efficacy (Khenfer, Roux, Tafani, & Laurin, 2016b) or when they think about the detailed list of actions required (Khenfer, Laurin, Tafani, Roux, & Kay, 2016a)—but not when they are made to feel confident, powerful Gods restore their motivation to pursue their goals.

To summarize, the by-products of God's powerful nature when it comes to self-regulation are multifaceted. Powerful Gods can either bolster or undermine individual goal pursuit motivation, and while early findings point to one possible way of predicting which of these effects will emerge, future research may further confirm this idea or propose others.

I turn now to the second, related domain of powerful Gods' by-products.

Prosocial Punishment

Another consequence of the idea of God's power, perhaps in conjunction with the idea of God's moral concern, relates to people's willingness to take on the particular social responsibility of punishing those who fail to help the group. This social responsibility, known as prosocial punishment, means going out of one's way, or incurring costs, in order to punish those who behave antisocially to dissuade them and others from doing so in the future. There is ample evidence, ranging from lab experiments to field investigations, that prosocial punishment helps societies remain cooperative (e.g., Barclay, 2006; Fehr & Gächter, 2000, 2002; Henrich et al., 2010). Moreover, those same studies show that people are willing to engage in this costly punishment even when they themselves do not stand to benefit directly from doing so. In other words, when people see Person A harm, or fail to cooperate, with Person B, even if both of those people are complete strangers, people are willing to spend their own resources to punish Person A.

But recall the notion of social loafing I alluded to earlier: People work less hard toward achieving a given end when they know they are not alone in the endeavor. The same principle applies here: People are less willing to engage in costly punishment when they know others have the opportunity to do so (O'Gorman, Henrich, & Van Vugt, 2008). Thus, if people view God as caring about human beings' moral conduct, and as powerful enough to punish those who fail to exhibit this moral conduct, they may loaf: They may choose to keep their resources for other purposes, and let God take on the responsibility of punishing wrongdoers. Consistent with this idea, at least one set of studies has found that reminders of God, particularly among people who believe in a powerful, interventionist God, make people less willing to spend their own money to punish another participant for failing to share resources appropriately with another person (Laurin et al., 2012b). Moreover, that same paper reports that these reminders make people less willing to see their tax dollars invested in efforts to catch and punish criminals. Interestingly, however, these same studies also find that these effects run counter to the effects of religion

more broadly: That is, being religious in general makes people *more* willing to invest their own resources in punishing wrongdoers, while the specific belief in a powerful, interventionist God makes them *less* willing to do so.

This same idea of outsourcing costly punishment to God may apply to other domains of social responsibility as well (but see Be'ery & Bloom, 2015). That is, whenever people see that God's likely interventions align with areas of social responsibility, they may choose to let God take charge. For instance, if people believe that God cares to preserve the environment, they may decide to leave sustainability to God, and save themselves the trouble of composting, or carrying their own grocery bags, or walking rather than driving (see also Meijers & Rutjens, 2014). To compound this issue further, people have an aversion to playing God: They are uncomfortable with human beings playing roles that are supposed to be God's (Waytz & Young, 2016). That is, to the extent people see preserving the greater good as God's domain, they may be not only selfishly unwilling to help out, but also cognitively uncomfortable with *any* human beings doing so.

To summarize, God's powerful nature may produce mixed effects on people's willingness to invest efforts in their own pursuits, while also diminishing people's willingness to engage in prosocial punishment. This last finding raises broader questions about social responsibility; future research may help add nuance on this point.

By-products of God's MORALLY CONCERNED Nature

Third, and finally, what might be the by-products of the fact that Big Gods are represented as being concerned with human morality? The previous section alluded to this feature of Big Gods: No research has yet demonstrated this, but the effect whereby reminders of powerful Gods reduce people's willingness to take on prosocial punishment may require that these powerful Gods also be invested in human morality. The logic underlying that hypothesis relies on the idea that Big Gods' morally invested nature makes them want to punish humans who do wrong. In this section, I focus on the other side of that same coin: The idea that Big Gods' morally invested nature makes them want to reward—or at the very least protect—humans who do right.

The cultural evolution hypothesis emphasizes Big Gods' ability to punish people who do wrong; that is a key mechanism through which Big Gods can serve the adaptive function of enforcing cooperation. But it is almost impossible to think about an entity who punishes wrongdoers without rewarding, in some fashion, good-doers. At the very least, by not punishing those who do good, Big Gods would be ensuring better outcomes for them, compared to those who do wrong. But it seems likely that even beyond that, people have difficulty conceiving of a God who is *only* punitive toward the sinful, without also assuming that that same God is also benevolent toward the virtuous (see e.g., Lawrence, 1997; Shariff & Norenzayan, 2011; Shariff & Rhemtulla, 2012).

If people assume that God rewards or protects the virtuous, what does that mean for people's assumptions about how God will treat *them*? It is one of the most fundamental laws of psychology that most people think they are pretty great (e.g., Alicke, 1985; Brown, 1986; Dunning, Heath, & Suls, 2004). This is true on all sorts of dimensions, including morality (Brown, 2007): A recent study showed that even incarcerated criminals, including those imprisoned for violence against others, rate themselves as more moral than most other prisoners, and more moral than most other community members (Sedikides, Meek, Alicke, & Taylor, 2014).

So it is assumed that people believe that God rewards or protects those who are virtuous, and they believe that they are more virtuous than most other people. It therefore stands to reason that most people believe that God will reward or protect *them*. Supporting this idea, recent findings indicate that when people think about God prior to deciding whether to take a risk, they feel protected, and therefore become more likely to take that risk (Kupor, Laurin, & Levav, 2015). For instance, participants who first read a passage about God, compared to those who read an irrelevant passage, said they were more likely to actually take a risk they had been considering, expressed more interest in engaging in the risky activity of skydiving, and were more likely to choose an experimental task that could potentially harm their eyesight, compared to a safer task. Moreover, meditational analyses indicated that the reason God reminders made people more interested in taking risk was that God reminders made people feel they were less likely to suffer harm from doing so.

Two additional pieces of data from this same paper add nuance to these findings. First, reminders of God do *not* make people more likely to take a risk that has immoral connotations (e.g., attempt to bribe someone with a slight risk of being discovered), which makes sense: God protects the virtuous, not the sinful. And second, when the divine protection that people anticipate fails to materialize, people feel angry with God: When participants are induced to take a very small risk which, unbeknownst to them, is rigged such that they will always fail, those previously primed with God report more negative feelings toward God than participants in a no-prime control condition. This both further confirms the idea that people did, in fact, expect that God would protect them (what other reason would they have to feel angry?), and carries significant implications: Negative affect toward God is psychologically maladaptive, in that it is linked with increased anxiety, depressive affect, poor coping, and, in the long run, even increased mortality (Exline & Rose, 2005; Exline, Yali, & Lobel, 1999; Pargament, Koenig, Tarakeshwar, & Hahn, 2001; Pargament et al., 1998).

Another finding that describes a by-product of God's moral concern, and more specifically of people's belief that God will protect the virtuous and therefore protect them, comes from work describing God as a relationship partner (e.g., Granqvist, Mikulincer, & Shaver, 2010; Kirkpatrick, 1998; Pollner, 1989). If I believe that God protects the virtuous, and therefore protects me, it is no great leap for me to then believe that God cares about me, and perhaps even that God likes or loves me. If people feel loved and cared for by God, then this may mean that God can be a source of the same kinds of benefits people get from their relationships with other people.

One way of illustrating that idea comes from a recent paper showing that, when their interpersonal closeness with a romantic partner is threatened, people will seek to compensate for that threat by drawing closer in their relationship with God. Conversely, when people feel their closeness with God is threatened, they compensate for *that* threat by drawing closer to their relationship partners (Laurin, Schumann, & Holmes, 2014). Moreover, this occurs *only* among people with high self-esteem, whose default response to potentially losing interpersonal closeness is to seek to reestablish that closeness, not withdraw into themselves (Murray, Derrick, Leder, & Holmes, 2008; Murray, Holmes, & Collins, 2006).

To summarize, most people assume that God will reward or protect them, perhaps because most people assume that they are better, morally, than others, and that therefore a morally concerned God will keep them from harm. Perhaps as a result of these beliefs, the idea of God can make people take risks they otherwise might not, so long as those risks do not imply any kind of immorality. Moreover, they may view God as a relationship partner capable of providing benefits that are interchangeable with the benefits they receive from their human relationship partners.

By-products: A Summary

In short, the three characteristics that define Big Gods generate a number of by-product effects that occur alongside their hypothesized original adaptive function. Because they are watchful, Big Gods can help people self-monitor better, and promote their willingness to withstand temptation, both of which make for better self-regulation. Because they are powerful, Big Gods can motivate people to invest efforts in pursuit of their goals, by reminding them that the world is structured, orderly and sensible, but can also demotivate them if people interpret the Big Gods as interfering with or substituting for their ability to control their own success. Moreover, Big Gods who are both powerful *and* morally invested can detract from people's willingness to take responsibility for punishing wrongdoers. Finally, because they are morally concerned, and therefore likely to reward the virtuous, most people likely view Big Gods as a source of protection and interpersonal closeness.

Conclusion: Ongoing Cultural Evolution of God Beliefs

We often make the mistake of thinking of evolution in historical terms, implying, perhaps unintentionally, that it is a process that has reached its final destination. This is not so. In this chapter I have both outlined one prominent theory of the origin of belief in Big Gods—that they emerged through cultural evolution because they promote large-scale cooperation—and described a set of findings regarding the by-products, both positive and negative, of this evolved belief. One important task

for future research, then, is to unpack how all of these by-products contribute to the net culturally adaptive consequences of Big God beliefs, which should enable us to predict the trajectory of those beliefs at the population level. That is, we must understand how the by-products either add to or subtract from the original adaptive function of Big God beliefs, so that we can determine whether their overall effect helps cultures survive, or instead leads to their demise.

Achieving this understanding requires first that we quantify the net effect of each of the by-products of belief in Big Gods: To what extent do these by-products make a group or society more versus less likely to flourish? In some instances the answers seem at least directionally clear. Will a society full of individuals who are better self-monitors and better able to resist temptation thrive more than a society full of individuals who are oblivious to themselves and unable to restrain their basest impulses? Probably. But will a society full of risk-seeking individuals who are unwilling to take on the responsibility of prosocial punishment flourish more than a society full of cautious enforcers of prosociality? Probably not. But critically, researchers must establish the magnitude of these effects, in order to determine the net impact of all the by-products combined.

However, this is only the first step. To predict the ongoing evolution of Big God beliefs, we must also quantify the extent to which the cooperation-promoting function of Big Gods continues to be something cultural evolution will select for. Some evidence indicates that civic institutions can implement the same kinds of surveillance and punishment, and thereby promote the same kind of cooperation, as Big Gods (e.g., Shariff & Norenzayan, 2007). In other words, Big God beliefs may have lost their edge in terms of why they originally propagated, and their presence may no longer play such a crucial role in ensuring large-scale cooperation. Quantifiable answers to the questions of how much Big God beliefs help ensure human cooperation, and of how much their side effects promote or prevent societal flourishing, will allow us to predict how Big God beliefs will continue to evolve, and explain how their prevalence has shifted over the past centuries.

References

- Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, 49(6), 1621–1630.
- Arkin, R. M., & Baumgardner, A. H. (1986). Self-presentation and self-evaluation: Processes of self-control and social control. In R. M. Baumeister (Ed.), *Public self and private self* (pp. 75–97). New York, NY: Springer.
- Atran, S., & Norenzayan, A. (2004). Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences*, 27(6), 713–770.
- Baldwin, M. W., & Holmes, J. G. (1987). Salient private audiences and awareness of the self. *Journal of Personality and Social Psychology*, 52(6), 1087–1098.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65–94.

- Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, 27(5), 325–344.
- Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Science*, 4(1), 29–34.
- Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a non-natural entity: Anthropomorphism in God concepts. *Cognitive Psychology*, 31(3), 219–247.
- Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, 91(1), 3–26.
- Be'ery, G., & Bloom, P. B.-N. (2015). God and the welfare state—substitutes or complements? An experimental test of the effect of belief in God's control. *PLoS ONE*, 10(6), e1238858.
- Bock, W. J. (2000). Explanatory history of the origin of feathers. *American Zoologist*, 40(4), 478–485.
- Botero, C. A., Gardner, B., Kirby, K. R., Bulbulia, J., Gavin, M. C., & Gray, R. D. (2014). The ecology of religious beliefs. *PNAS*, 111(47), 16784–16789.
- Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York, NY: Basic Books.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, 132(3), 337–356.
- Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, 4(4), 353–376.
- Brown, J. D. (2007). *The self*. New York, NY: Psychology Press.
- Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, 53(5), 533–548.
- Carter, E. C., McCullough, M. E., & Carver, C. S. (2012). The mediating role of monitoring in the association of religion with self-control. *Social Psychological and Personality Science*, 3(6), 691–697.
- Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press.
- Carver, C. S., & Scheier, M. F. (2011). Self-regulation of action and affect. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (2nd ed., pp. 3–21). New York, NY: Guilford.
- Cavalli-Sforza, L., & Feldman, M. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton, NJ: Princeton University Press.
- Chae, B. G., & Zhu, R. J. (2014). Environmental disorder leads to self-regulatory failure. *Journal of Consumer Research*, 40(6), 1203–1218.
- Chudek, M., Brosseau-Liard, P. E., Birch, S., & Henrich, J. (2013). Culture-gene coevolutionary theory and children's selective social learning. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 181–185). New York, NY: Oxford University Press.
- Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, 15(5), 218–226.
- Dahl, D. W., Manchanda, R. V., & Argo, J. J. (2001). Embarrassment in consumer purchase: The roles of social presence and purchase familiarity. *Journal of Consumer Research*, 28(3), 473–481.
- Darwin, C. (1877/2004). *The descent of man, second edition, with an introduction by Adrian Desmond and James Moore*. London: Penguin.
- de Waal, F. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59(1), 279–300.
- Dijksterhuis, A., Preston, J., Wegner, D. M., & Aarts, H. (2008). Effects of subliminal priming of self and God on self-attribution of authorship for events. *Journal of Experimental Social Psychology*, 44(1), 2–9.
- Dunbar, R. I. M. (2003). Evolution of the social brain. *Science*, 302(5648), 1160–1161.

- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69–106.
- Eitam, B., & Higgins, E. T. (2010). Motivation in mental accessibility: Relevance of a representation (ROAR) as a new framework. *Social and Personality Psychology Compass*, 4(10), 1–17.
- Evans, E. M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, 42(3), 217–266.
- Exline, J. J., & Rose, E. (2005). Religious and spiritual struggles. In R. F. Paloutzian & C. L. Park (Eds.), *Handbook of the psychology of religion and spirituality* (pp. 315–330). New York, NY: Guilford Press.
- Exline, J. J., Yali, A. M., & Lobel, M. (1999). When God disappoints: Difficulty forgiving God and its role in negative emotion. *Journal of Health Psychology*, 4(3), 365–379.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4), 980–994.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140.
- Gervais, W. M., & Norenzayan, A. (2012). Like a camera in the sky? Thinking about god increases public self-awareness and socially desirable responding. *Journal of Experimental Social Psychology*, 48(1), 298–302.
- Granqvist, P., Mikulincer, M., & Shaver, P. R. (2010). Religion as attachment: Normative processes and individual differences. *Personality and Social Psychology Review*, 14(1), 49–59.
- Gray, K., & Wegner, D. M. (2010). Blaming god for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review*, 14(1), 7–16.
- Greenspan, S. (2006). *How symbols, language, and intelligence evolved from early primates to modern humans*. Cambridge, MA: Da Capo Press.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327(5972), 1480–1484.
- Henrich, N., & Henrich, J. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. Oxford: Oxford University Press.
- Hitsch, G. J., Hortacsu, A., & Ariely, D. (2010). Matching and sorting in online dating. *American Economic Review*, 100(1), 130–163.
- James, W. (1880). Great men, great thoughts, and the environment. *Atlantic Monthly*, 46(276), 441–459.
- Kanfer, F. H., & Karoly, P. (1972). Self-control: A behavioristic excursion into the lion's den. *Behavioral Theory*, 3(3), 398–416.
- Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65(4), 681–706.
- Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the Government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology*, 95(1), 18–35.
- Kay, A. C., Laurin, K., Fitzsimons, G. M., & Landau, M. J. (2014). A functional basis for structure-seeking: Exposure to structure promotes willingness to engage in motivated action. *Journal of Experimental Psychology: General*, 143(2), 486–491.
- Khenfer, J., Laurin, K., Tafani, E., Roux, E., & Kay, A. C. (2016a). Here's what you gotta do! Reminders of external agency affect motivational responses to savings advice in opposing directions. Manuscript under review.
- Khenfer, J., Tafani, E., Roux, E., & Laurin, K., (2016b). When God's (not) needed: The divergent influence of belief in divine control on goal commitment. Manuscript under review.
- Kim-Spoon, J., Farley, J. P., Holmes, C., Longo, G. S., & McCullough, M. E. (2014). Processes linking parents' and adolescents' religiousness and adolescent substance use: Monitoring and self-control. *Journal of Youth and Adolescence*, 43(5), 745–756.
- Kirkpatrick, L. A. (1998). God as a substitute attachment figure: A longitudinal study of adult attachment style and religious change in college students. *Personality and Social Psychology Bulletin*, 24(9), 961–973.

- Kitchens, M. B. (2015). Thinking about God causes internal reflection in believers and unbelievers of God. *Self and Identity*, 14(6), 724–747.
- Kupor, D. M., Laurin, K., & Levav, J. (2015). Anticipating divine protection? Reminders of God can increase nonmoral risk taking. *Psychological Science*, 26(4), 374–384.
- Landau, M. J., Kay, A. C., & Whitson, J. (2015). Compensatory control and the appeal of a structured world. *Psychological Bulletin*, 141(3), 694–722.
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, 36(4), 343–356.
- Laurin, K., Fitzsimons, G. M., & Kay, A. C. (2011). Social disadvantage and the self-regulatory function of justice beliefs. *Journal of Personality and Social Psychology*, 100(1), 149–171.
- Laurin, K., & Kay, A. C. (2016). Religion and self-regulation: Integrating skills-based and motivation-based accounts. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, & applications* (pp 305–322). Guilford Publications.
- Laurin, K., Kay, A. C., & Fitzsimons, G. M. (2012a). Divergent effects of activating thoughts of God on self-regulation. *Journal of Personality and Social Psychology*, 102(1), 4–21.
- Laurin, K., Schumann, K., & Holmes, J. G. (2014). A relationship with God? Connecting with the divine to assuage fears of interpersonal rejection. *Social Psychological and Personality Science*, 5(7), 777–785.
- Laurin, K., Shariff, A. F., Henrich, J., & Kay, A. C. (2012b). Outsourcing punishment to God: Beliefs in divine control reduce earthly punishment. *Proceedings of the Royal Society B*, 279 (1741), 3272–3281.
- Lawrence, R. T. (1997). Measuring the image of God: The God image inventory and the God image scales. *Journal of Psychology & Theology*, 25(2), 214–226.
- Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior*. Boulder, CO: Westview.
- Lieberman, P. (1991). *Uniquely human: The evolution of speech, thought, and selfless behavior*. MA: Harvard University Press.
- Lumsden, C., & Wilson, E. O. (1981). *Genes, mind and culture: The coevolutionary process*. Cambridge, MA: Harvard University Press.
- McCullough, M. E., & Willoughby, B. L. B. (2009). Religion, self-regulation, and self-control: Associations, explanations, and implications. *Psychological Bulletin*, 135(1), 69–93.
- McNamara, R. A., Norenzayan, A., & Henrich, J. (2016). Supernatural punishment, in-group biases, and material insecurity: Experiments and ethnography from Yasawa, Fiji. *Religion, Brain, and Behavior*, 6(1), 34–55.
- Meijers, M. H. C., & Rutjens, B. T. (2014). Affirming belief in scientific progress reduces environmentally friendly behavior. *European Journal of Social Psychology*, 44(5), 487–495.
- Mikalson, J. (2010). *Ancient Greek religion*. Blackwell Publishing.
- Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, 126(2), 247–259.
- Murray, S. L., Derrick, J. L., Leder, S., & Holmes, J. G. (2008). Balancing connectedness and self-protection goals in close relationships: A levels-of-processing perspective on risk regulation. *Journal of Personality and Social Psychology*, 94(3), 429–459.
- Murray, S. L., Holmes, J. G., & Collins, N. L. (2006). Optimizing assurance: The risk regulation system in relationships. *Psychological Bulletin*, 132(5), 641–666.
- Norenzayan, A. (2013). *Big Gods: How religion transformed cooperation and conflict*. Princeton, NJ: Princeton University Press.
- Norenzayan, A., & Atran, S. (2004). Cognitive and emotional processes in the cultural transmission of natural and nonnatural beliefs. In M. Schaller & C. Crandall (Eds.), *The psychological foundations of culture* (pp. 149–169). Hillsdale, NJ: Lawrence Erlbaum Associates.
- O’Gorman, R., Henrich, J., & Van Vugt, M. (2008). Constraining free riding in public good games: Designated solitary punishers can sustain human cooperation. *Proceedings of the Royal Society B*, 276(1655), 323–329.
- Panchanathan, K., & Boyd, R. (2003). A tale of two defectors: The importance of standing for the evolution of indirect reciprocity. *Journal of Theoretical Biology*, 224(1), 115–126.

- Pargament, K. I., Koenig, H. G., Tarakeshwar, N., & Hahn, J. (2001). Religious struggle as a predictor of mortality among medically ill elderly patients: A 2-year longitudinal study. *Archives of Internal Medicine*, *161*(15), 1881–1885.
- Pargament, K. I., Zinnbauer, B. J., Scott, A. B., Butter, E. M., Zerowin, J., & Stanik, P. (1998). Red flags and religious coping: Identifying some religious warning signs among people in crisis. *Journal of Clinical Psychology*, *54*(1), 77–89.
- Peoples, H. C., & Marlowe, F. W. (2012). Subsistence and the evolution of religion. *Human Nature*, *23*(3), 253–269.
- Pew Research Center. (2012). *Religiously unaffiliated*. Retrieved from <http://www.pewforum.org/2012/12/18/global-religious-landscape-unaffiliated/>
- Pollner, M. (1989). Divine relations, social relations, and well-being. *Journal of Health and Social Behavior*, *30*(1), 92–104.
- Preston, J. L., & Ritter, R. S. (2013). Different effects of religion and God on prosociality with the ingroup and the outgroup. *Personality and Social Psychology Bulletin*, *39*(11), 1471–1483.
- Purzycki, B. G., Finkel, D. N., Shaver, J., Wales, N., Cohen, A. B., & Sosis, R. (2012). What does God know? Supernatural agents' access to socially strategic and nonstrategic information. *Cognitive Science*, *36*(5), 846–869.
- Pyytiäinen, I. (2009). *Supernatural agents: Why we believe in souls, Gods, and Buddhas*. New York, NY: Oxford University Press.
- Randolph-Seng, B., & Nielsen, M. E. (2007). Honesty: One effect of primed religious representations. *The International Journal for the Psychology of Religion*, *17*(4), 303–315.
- Rives, J. (2007). *Religion in the Roman Empire*. Blackwell Publishing.
- Roes, F. L., & Raymond, M. (2003). Belief in moralizing gods. *Evolution and Human Behavior*, *24*(2), 126–135.
- Saroglou, V. (2011). Believing, bonding, behaving, and belonging: The big four religious dimensions and cultural variation. *Journal of Cross-Cultural Psychology*, *42*(8), 1320–1340.
- Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Monterey, CA: Brooks-Cole.
- Schlenker, B. R. (1986). Self-identification: Toward an integration of the private and public self (pp. 21–62). In R. F. Baumeister (Ed.), *Public self and private self*. New York, NY: Springer.
- Schvaneveldt, R. W., & Meyer, D. E. (1973). Retrieval and comparison processes in semantic memory (pp. 395–409). In S. Kornblum (Ed.), *Attention and performance IV*. New York, NY: Academic Press.
- Sedikides, C., Meek, R., Alicke, M. D., & Taylor, S. (2014). Behind bars but above the bar: Prisoners consider themselves more prosocial than non-prisoners. *British Journal of Social Psychology*, *53*(2), 396–403.
- Shariff, A. F., & Norenzayan, A. (2007). God is watching you: Priming god concepts increases prosocial behavior in an anonymous economic game. *Psychological Science*, *18*(9), 803–809.
- Shariff, A. F., & Norenzayan, A. (2011). Mean gods make good people: Different views of god predict cheating behavior. *International Journal for the Psychology of Religion*, *21*(2), 85–96.
- Shariff, A. F., & Rhemtulla, M. (2012). Divergent effects of beliefs in heaven and hell on national crime rates. *PLoS ONE*, *7*(6), e39048.
- Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2016). Religious priming: A meta-analysis with a focus on religious prosociality. *Personality and Social Psychology Review*, *20*(1), 27–48.
- Slingerland, E. (2013). Body and mind in early China: An integrated humanities-science approach. *Journal of the American Academy of Religion*, *81*(1), 6–55.
- Smith, R. J. (1996). Biology and body size in human evolution: Statistical inference misapplied. *Current Anthropology*, *37*(3), 451–481.
- Sumida, S. S., & Brochu, C. A. (2000). Phylogenetic context for the origin of feathers. *American Zoologist*, *40*(4), 486–503.
- Vail, K. E., III, Adndt, J., & Abdollahi, A. (2012). Exploring the existential function of religion and supernatural agent beliefs among Christians, Muslims, atheists, and agnostics. *Personality and Social Psychology Bulletin*, *38*(10), 1288–1300.

- Vroom, V. H. (1964). *Work and motivation*. San Francisco, CA: Jossey-Bass.
- Waytz, A., & Young, L. (2016). Aversion to playing God: A key principle underlying moral judgment. Unpublished manuscript.
- Williams, T. N., Mwangi, T. W., Roberts, D. J., Alexander, N. D., Weatherall, D. J., Wambua, S., et al. (2005). An immune basis for malaria protection by the sickle cell trait. *PLoS Medicine*, 2(5), e128.
- Yilmaz, O., & Bağçekapili, H. G. (2016). Supernatural and secular monitors promote human cooperation only if they remind of punishment. *Evolution and Human Behavior*, 37(1), 79–84.
- Zhong, C.-B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science*, 21(3), 311–314.

From the Impossible to the Improbable: A Probabilistic Account of Magical Beliefs and Practices Across Development and Cultures

Martin Fortier and Sunae Kim

Like scientists, children and lay people are eager to build theories in order to make sense of their surrounding environment (e.g., Furnham, 1988). The scholars studying such lay theories like to draw suggestive parallels between scientific and lay theories: although the latter and the former have some undeniable dissimilarities, it is often remarked that they also share a great deal of commonalities (for a discussion, see: Gopnik, Wellman, & Kuhl, 1999; Kuhn, 1989). Interestingly, the domain of magic seems to provide a nice counterexample. Lay people uniquely develop complex theories—with no significant scientific counterparts—aiming to explain special features of “supernatural” events and beings.

Previous studies have demonstrated that lay people often develop distinct competing theories to explain the world. These lay theories are variably distributed across individuals, contexts, occupations, social groups, and cultures (Dweck, 2006; Markus & Hamedani, 2007). This chapter focuses on the “building blocks” of lay theories of magic. Existing studies on supernatural thinking maintain that lay theories of magic stem from a single common mechanism triggered by impossible events. In contrast to this widespread view, in this chapter, we will argue that lay theories of magic stem from at least two very distinct mechanisms: one triggered by impossible events and another triggered by improbable ones. We will show that, as

M. Fortier (✉)

Ecole Normale Supérieure, Ecoles des Hautes Etudes en Sciences Sociales,
Institut Jean Nicod, PSL Research University, Paris, France
e-mail: martin.fortier@ens.fr

S. Kim

Institute of Psychology, Department of Developmental and Clinical Child Psychology,
Eötvös Loránd University, Budapest, Hungary

yet, the diversity of the “building blocks” of lay theories of magic, and, as a consequence, the diversity of these theories themselves, has been largely underestimated.

In order to better grasp the originality of our proposal, it will be useful to briefly introduce existing approaches to magic. Magical thinking has been extensively studied by developmental psychologists and anthropologists alike, but seldom have these two approaches been discussed together. This chapter precisely tries to fill this gap by bringing together recent findings in developmental psychology, cognitive psychology, cognitive science of religion, and anthropology. Our central claim is that psychological and anthropological studies account for magic in two very different ways. Psychological evidence shows that people resort to magical explanations when faced with *impossible* events whereas anthropological evidence suggests that magical explanations are usually triggered by *improbable* events (rather than impossible ones). In this chapter, we argue that these two approaches to magic are not two partial accounts of the same cognitive phenomenon but two independent lines of research concerning distinct cognitive mechanisms. If, as we think, two distinct concepts of magic are to be recognized—one being based on counterintuition and the other on probabilistic reasoning—then an important question arises: how are these two concepts of magic, respectively, distributed across development and cultures?

The first part of the chapter focuses on the psychological approach to magic. These studies show that starting at a young age children distinguish between events and entities that violate our intuitive notions of basic causal laws (e.g., gravity) and those that do not. Next, we introduce the anthropological debate about the boundaries of magic (how it differs from science and religion), and we examine a series of case studies (on witchcraft and shamanism) which nicely epitomize the anthropological understanding of magic. Drawing upon ethnographic evidence, we argue that the approach to magic in terms of impossible or counterintuitive events fails to account for people’s actual beliefs and practices. Part three is entirely dedicated to the formalization of anthropological data discussed in part two. We argue that algorithmic complexity provides the best tools to model the way magic—as described by anthropologists—works; it is thus proposed that magical explanations are typically triggered by *complexity drops*, i.e., by events and objects whose features are expected to be complex but turn out to be simple. Part four attempts to provide an answer as to how “counterintuitive-magic” and “probabilistic-magic” are distributed across development and cultures. We first explore the historical trajectory of magic by tackling the themes of secularization and explanatory coexistence. Next, we explain why the anthropological account of magic poses a real challenge to psychological approaches to magic. Finally, we put forward four hypotheses specifying how “counterintuitive-magic” and “probabilistic-magic” are likely to be combined across development and distributed across cultures; we examine the strengths and weaknesses of each of these hypotheses and discuss the future directions of the study of magic.

The Developmental Psychological Approach: The Counterintuitiveness of Magic

Children's Conception of Magic

A large number of empirical studies show that young children are able to differentiate between impossible and possible events (Johnson & Harris, 1994; Rosengren, Kalish, Hickling, & Gelman, 1994). These impossible events (e.g., a person flying in the air) contrary to possible events, involve some sorts of violations of natural physical laws. Thus, children's distinction between impossible and possible events demonstrates their understanding of ordinary causality, or how the world works. Moreover, young children understand that some events involve ordinary processes whereas others involve magical processes. The theoretical claim is that children evoke a notion of magic to explain the events that cannot be explained by ordinary forces or processes (Chandler & Lalonde, 1994; Phelps & Woolley, 1994). By implication, children's understanding of magic is intimately related to their understanding of ordinary and possible events.

Young children have a very strict sense of what is impossible and possible either due to their less advanced understanding of causality or lack of familiarity or other reasons (see below, "Factors in children's judgment of reality vs. magical status"). Shtulman and Carey (2007) demonstrated that children below the age of 8 categorized improbable events, those that are possible but highly unlikely to happen in the real world (e.g., "polka dots on an airplane"), as impossible. From this finding, they argued that children's understanding of a distinction between impossible and possible events, and of ordinary causality, still undergoes a development. Even when a plausible explanation of how an improbable event can happen was offered, children continued to treat it as impossible (Woolley & Ghossainy, 2013). However, it is not clear whether children think of these improbable events as requiring a magical process or something different. Children have some understanding that different types of constraints are involved in the violation of different types of laws (Kalish, 1998). For example, they understand violations in the domain of social norms (e.g., violating school rules) are different from those in the domain of physical laws (e.g., walking through the wall). Given this consideration, children in Shtulman and Carey's (2007) study could think that improbable events are impossible for a different reason; they may think that these events are more similar to those involving violations of regularity in their social experiences (e.g., "dots on an airplane is not possible because it is not how it is supposed to be") than violations of physical laws.

Children's understanding of 'magic' changes with age so that 4-year-old children tend to think of magic as involving special power or skills, whereas 5-year-old children tend to think of it as involving tricks (Rosengren & Hickling, 1994). Nevertheless, magical and non-magical thinking coexist within an individual

throughout development (Subbotsky, 1993, 2010). For example, children who are initially skeptical about magical events can become believers. In one study, for instance, children were skeptical about a magical process of turning a drawing into a real entity that was represented in the drawing; but after being told about a story of a girl who used a box to turn drawings into real entities, a majority of the children when left alone displayed performances that reflected their magical stance toward the box (Subbotsky, 1993). In another study, while children initially denied the reality status of an event that was verbally described, they became credulous about the event once the event was displayed live to them (Woolley, Boerger, & Markman, 2004). Moreover, children who were led to believe a supernatural being communicated to them via signs (e.g., a light flickering) were more prone to the communicative message than those who were not (Bering & Parker, 2006). In the study, half of the children aged 3–9 years were assigned to an experimental group and, the other half, a control group. Children in the experimental group were introduced to an invisible agent called Princess Alice who they were told would help them choose a correct answer option in a game. These children were not told explicitly how Princess Alice would help but simply instructed somehow they would be told. Then, whenever they made an inaccurate response choice, a light flickered, for example. Children in the control group did not receive such priming of the idea about an invisible agent. The children who were led to believe in the existence of an invisible agent were more likely than those who were not believed that the invisible agent communicated to them in order to help them with the game. Interestingly, the effect was observed among children older than 7 years old, not among younger children. This age difference may be because attribution of the communicative intention as well as epistemic states to the supernatural agents requires further theory of mind development, and an understanding of the symbolic nature of the communicative message may not be present in younger children. Finally, contrary to a theoretical view that with age and education magical thinking declines and is replaced by scientific thinking (e.g., Piaget, 1929), adults do not completely outgrow magical thinking. Subbotsky (1993) demonstrated that adults were quite credulous to the idea that an experimenter was capable of performing impossible events (e.g., making an object disappear) after they viewed the impossible events. He further demonstrated that a scientific mode of thinking did not replace magical thinking even with increasing age, but the two coexisted in people's minds; and magical orientation was observed in behaviors even when verbal responses reflect a scientific mode of thinking (Subbotsky, 2001). For example, seeing an object destroyed after being placed inside an empty box without any physical and spatial contact with the object, but accompanied by a magical spell, participants tended to deny in their verbal judgments the possibility that the magical spell had caused the event, and yet they refused to place another object inside the box when asked to.

Factors in Children's Judgment of Reality Versus Magical Status

Of course, whether and to what extent children treat a certain event as magical depends on various factors (see Woolley & Ghossainy, 2013, for a review). Here we highlight two: familiarity with an event and social and cultural contexts. Children's familiarity with an event affects their judgment of reality status of an event. Cook and Sobel (2011) tested 4- and 6-year-old children and adults' assignment of reality status of possible and impossible machines. They found that children as well as adults distinguished between machines that violated biological or physical casual laws and those that did not, but 4-year-old children were more likely to judge familiar possible machines as real than unfamiliar possible machines. In addition, social and cultural contexts such as children's religious education and background can modulate their judgment of reality status of certain events. Four- to six-year-old children from religious families were more likely to judge religious context stories (e.g., biblical stories from the Old Testament) as real than children from non-religious families (Vaden & Woolley, 2011).

Next, in the following two sections, we focus our discussion on a handful of studies on children's attribution of magical power to themselves and others.

Children's Attribution of Magical Power to Themselves

Children's Wishing and Imagination

Young children believe that their own imagination can change or create reality. Woolley and Wellman (1993) demonstrated that, after imagining something in an empty box, 3-year-olds responded affirmatively that the box contained the object imagined. Harris, Brown, Marriot, Whittall, and Harmer (1991) also showed that 4- to 6-year-olds behaved as if the imagined object was real. In the study, children were asked to pretend that a monster or a bunny was inside a box. When they were left alone in the room with the box, children peeked into the bunny box but not into the monster box. However, there is also evidence that children acknowledged the existence of an imagined object in a box, but did not act accordingly, so that when another person asked for the imagined object, children rarely offered the box to the person (Phelps & Woolley, 1994). Nevertheless, it is possible that children in this study may have believed in the event, but did not outwardly show this to another person who may not have endorsed their magical world.

Children's belief that imagination can change the state or course of a real event may be extended to another case; wishing for something. Vikan and Clausen (1993) showed that children of ages 4 and 6 believed that wishing has an effect on another person, in terms of changing another person's actions. Moreover, children see wishing as a magical process involving special power or magic, not an ordinary

process involving mental-physical causality (Woolley, Phelps, Davis, & Mandell, 1999). These studies show that, with increasing age, children's tendency to believe in the efficacy of wishing wanes, and older compared to younger children thought that an event realized via wishing was caused by a trick rather than real magic (see also Woolley, 2000). A more recent study, however, shows that older children and adults preserve a tendency to believe in the efficacy of wishing (Subbotsky, 2005). Different methodologies used in the aforementioned studies make it difficult to adjudicate their conflicting findings. However, speculatively, they do suggest at the very least that magical tendency may be susceptible to contexts and conditions in which it is measured.

Children's Attribution of Magical Power to Other Agents

Children's Understanding of Supernatural Agents

Children believe in various supernatural and fantastical characters such as Santa Claus or the Tooth Fairy. All these characters share a common feature of possessing supernatural abilities violating physical and biological causal laws. We know that children treat supernatural agents differently from ordinary agents, attributing some special power only to the former. However, what exact capabilities children attribute to novel as well as familiar fantastical characters asks for further investigation. One line of research as reviewed below starts to provide some information concerning what those capabilities might be.

Children assign special power or ability to God. Vaden and Woolley (2011) showed that, while 4- and 6-year-old children viewed impossible events as not real, they endorsed the reality status of those events when they were told that God was involved. Vaden and Woolley asked 4- to 6-year-old children to judge the reality status of an event or a character described in a story either in a religious (reference to God, e.g., the story of Moses parting the sea but replacing original characters with novel character names) or a nonreligious context (otherwise the same stories as the religious context stories without reference to God). With increasing age, children who heard the religious context stories were more likely to judge the characters and events described as real than those who heard nonreligious stories.

Lane and colleagues, in a series of studies, extensively investigated what types of knowledge and abilities children and adults attribute to omniscient agents such as God as opposed ordinary human beings (Lane, Wellman, & Evans, 2010, 2012, 2014). Lane et al. (2012) demonstrated that children around the age of 5 attributed special power of knowing (e.g., knowing hidden contents of a box without perceptual access) to God and a novel agent, Mr. Smart, who was introduced as: ("This is Mr. Smart. Mr. Smart has special powers. He knows everything"). Lane et al. (2014) further investigated to what extent children and adults attributed extraordinary knowledge to special beings. With increasing age, participants tended to attribute broader and deeper knowledge to omniscient beings. A younger group

(3.5- to 6.5-year-olds) only attributed generic knowledge (e.g., “where to find the tallest tree in the world”) to Ms. Smart, whereas an older group (6.5- to 12-year-olds) attributed also their personal knowledge (e.g., birthdays) to the agent. Only adults, however, attributed the knowledge of their own thoughts and future to omniscient beings (Ms. Smart). Participants were also asked to choose between Ms. Smart (who was introduced as someone “who knows everything about everything”) and an expert (e.g., a mechanic) when asked who knows more about certain topics. With increasing age, participants tended to attribute deeper knowledge to omniscient beings so that, for example, only adults attributed more expertise-related knowledge to Ms. Smart than an expert, whereas the youngest group attributed more expertise-related knowledge to an expert than to Ms. Smart. Moreover, children’s exposure to religion was related to their attribution of broader knowledge to Ms. Smart—but this relation was only specific to the age range of 4–6. The authors speculated that this age range coincides with children’s developing theory of mind, and social cultural contexts (e.g., religions) may facilitate children’s understanding of extraordinary minds during this period. In addition, children from religious backgrounds appreciated extraordinary minds more readily than those from nonreligious backgrounds (Lane et al., 2012). These studies suggest that already by preschool age children are capable of appreciating the kinds of special knowledge that omniscient beings have, and with increasing age they become better able to reason about the omniscient beings’ extraordinary knowledge. As will be demonstrated below, this ability is not limited to children’s reasoning but influences their daily interactions and learning.

Children’s Preferential Learning from a Magical Person

Children also attribute magical power to an ordinary person if the person displays actions that are beyond ordinary causality (e.g., moving an object without touching it). Those children who perceive the person as possessing magical power tend to preferentially learn from that person. In a study by Kim and Harris (2014a), children of ages 3–8 were presented with video clips in which a person performs magical actions (e.g., moving a box without touching it) whereas another person performs non-magical actions (e.g., moving a box with her hand). Then, children heard these two people offering conflicting pieces of novel information (e.g., a label for a novel object, “a fep” vs. “a tog”) and were asked to endorse one of the information pieces. With increasing age children were less likely to prefer the information supplied by the magical person. Notably, however, individual difference rather than age better predicted preferential learning tendency; those children who were credulous to the reality status of the magical events tended to preferentially learn from the person, as compared to skeptical children. This tendency to preferentially learn from a magical person was not limited to the magical power of actions, but observed in the magical power of reading others’ minds (Kim & Harris, 2014b). Children preferentially learned from a person who accurately guessed others’ thoughts over another person who figured out others’ thoughts via communication.

In sum, children are able to distinguish impossible events from possible events, and at the same time, they do not entirely disregard magical thinking. In particular, with the right amount of evidence and supporting environment, adults as well as children entertain the magical world they themselves or others created. Children, moreover, preferentially learn everyday knowledge from a person they view as possessing magical power. Notably, one key element of children's understanding of magical events and supernatural agents and power in developmental psychology as reviewed above is their understanding of violations of basic physical laws. However, this notion of magic as defined as a blanket term for an event that a child's naïve theories cannot explain—while it has provided the field with fruitful directions for empirical investigations—excludes a prevalently common notion of magic in our everyday lives, as will be discussed in the following sections.

The Anthropological Approach: The Ordinariness of Magic

Introducing the Anthropological Approach to Magic

The methods used by anthropologists to study magic, and, as a consequence, the models of magic that have emerged from anthropological findings, are somewhat different from those to be found in developmental psychology. In the field of anthropology, the very concept of magic originates itself in discussions sparked by evolutionary anthropologists—notably Edward Tylor and James Frazer. These authors were primarily interested in understanding how magic—more broadly how “primitive” religions and cults (animism, totemism, witchcraft, etc.)—have gradually developed into more “sophisticated” forms (such as monotheism) and finally into science. Evolutionary anthropologists were thus willing to specify the psychological mechanisms leading from one evolutionary step to the other.

A few decades later, the evolutionary perspective largely gave way to investigations into the intricate coexistence of ordinary knowledge and magical thinking and the relations between magic, science, and religion (Evans-Pritchard, 1976; Hubert & Mauss, 1903; Malinowski, 1948). Magical thinking was not considered necessarily primitive and irrational anymore. This line of research has further been explored in the second half of the twentieth century (Horton, 1967a, b; Nader, 1996; Tambiah, 1990). In a similar vein, some scholars have begun to address more epistemological and philosophical questions relating magic to the theme of apparent irrationality and that of the incommensurability between worldviews (Bonney & Laugier, 2003; Da Costa, Bueno, & French, 1998; Hollis & Lukes, 1982; Jarvie & Agassi, 1967; Sperber, 1985; Triplett, 1994).

One important distinction between psychological studies of magic reviewed earlier in the chapter and those in the field of anthropology concerns methodologies. Classically, anthropologists do not resort to experiments in order to investigate a

topic. The forerunners of the discipline (such as Tylor or Frazer) were strictly armchair anthropologists speculating on the basis of missionaries' and travelers' outlandish reports; yet, at least from Bronisław Malinowski on, anthropologists have fruitfully resorted to participant observation: they have studied other people's thinking by living with these people and putting themselves in the same life settings, and looking at how these settings affect them (both cognitively and emotionally) (Goulet & Miller, 2007). Some anthropologists whose work on magic and witchcraft has been much vaunted have endorsed this method of participant observation in a particularly radical fashion and have thus revealed that even a modern and secular academic mind can easily be affected by magical beliefs and practices (Favret-Saada, 1980, Chap. 2, 1990; Luhrmann, 1991, Chap. 21).

While psychologists are mainly concerned with studying how children and adults perform specific tasks within experimental settings, anthropologists are first and foremost interested in documenting how magical thinking unfolds in natural settings. Typically, psychologists present impossible stimuli to participants and then look at what participants' responses are, whereas anthropologists typically look at daily practices and rituals and from there infer what is deemed magical in a given culture. The psychological method could be characterized as top-down: it assumes that the concept of magic is unambiguous and uniform and it operationalizes this concept through experiments. On the other hand, the anthropological method is best described as bottom-up: no assumption is made as to what magic amounts to and it is only the collection of ethnographic data through participant observation which determines what is deemed magical by people in real life. This methodological distinction will prove important in the next sessions.

Magic, Science, and Religion: Debating the Boundaries of Magic

A good amount of research in Anthropology is devoted to the study of similarities and differences between magic on one side and science and religion on the other side. Examining the boundaries of magic will help us better understand its scope and mechanisms. The position Malinowski (1948) took in this heated debate is particularly interesting. On the one hand, Malinowski maintains that the Melanesians—and more generally people living in any culture in which magic is pervasive—do not live in a world utterly different from ours, nor do they use a logic different from ours. They do not inhabit an ethereal and mystical world. The Melanesians described by Malinowski prove to be expert and ingenious gardeners and fishermen, and while reading *The Argonauts of the Western Pacific*, one realizes how adept they are at solving practical and rational problems. Thus, it clearly appears that Malinowski strongly disagrees with scholars such as Lucien Lévy-Bruhl (1923, 1927, 1966), who have somewhat exaggerated the mystical

nature of indigenous thinking. On the other hand, Malinowski demurs at the view—notably held by Frazer (1922)—according to which magic can be conflated with science. Admittedly, Malinowski grants that magic, very much like science, has very concrete and down-to-earth ends; but he argues that the function of magic remains very different from that of science and that of profane practical knowledge.

In *Magic, science and religion and other essays*, the middle way between the Lévy-Bruhlian view and the Frazerian view is delimited by a central dichotomy: the *sacred* and the *profane*. Malinowski advances that “[i]n every primitive community [...] there have been found two clearly distinguishable domains, the Sacred and the Profane; in other words, the domain of Magic and Religion and that of Science” (1948, p. 1). When the Melanesians are gardening or fishing, their profane know-how is recruited; but when their extended knowledge meets some limit and when they realize that their practical knowledge is not sufficient to control a crucial parameter of gardening or fishing, they then resort to totally different kinds of tools: tools pertaining to the sacred domain. Magic is used to compensate the limited scope of daily and scientific knowledge. Malinowski evocatively illustrates this idea through the example of gardening: “there is a clear-cut division: there is first the well-known set of conditions, the natural course of growth, as well as the ordinary pests and dangers to be warded off by fencing and weeding. On the other hand, there is the domain of the unaccountable and adverse influences, as well as the great unearned increment of fortunate coincidence. The first conditions are coped with by knowledge and work, the second by magic” (1948, p. 12). By the same token, Malinowski points out that no magic is used when fishing is done in the inner lagoon whereas magical procedures are widely used in open-sea fishing (1948, p. 14). As it happens, the first type of fishing involves almost no uncertainty while this other type of fishing is replete with uncertainty. Very much like science and practical knowledge, magic targets practical ends (e.g., improving the outcome of fishing or the crop); but unlike science and daily know-how, magic governs a specific plane of reality—the unpredictable parameters targeted by magical rituals that cannot be targeted by profane knowledge whatsoever.

The domain of magic (of the sacred), we are told, cannot be conflated with that of science (of the profane). However, it is interesting to note that this dualism is somewhat challenged by some ethnographic findings. For example, Shipibo¹ hunters of the Peruvian Amazon have at their disposal an argosy of magical techniques to ensure luck and success in the forest. Among them are plant baths. Specific plants (e.g., *ochiti jana*) are purported to imbue one’s body with some magical smell or wind (*nihue*) thanks to which hunting expeditions will prove to be remarkably successful. This beneficial effect is taken to be mediated by the spirit (*yoshin*) of the plant used in the bath. Interestingly, this spirit is conceived as benevolently interfering in the hunting process both as an invisible intentional entity and as a perceptible smell. In such a case, it is notoriously difficult to

¹Data about Shipibo culture have been collected by one of us (MF) on his fieldwork located in the Peruvian Amazon. Magic in Shipibo culture will be further discussed below.

disentangle ordinary knowledge from magical stratagem. The smell of the plant obviously conceals the hunter's characteristic smell (which would otherwise frighten and drive away the game animals) and it also serves as an enticement (because, it seems, the game animals like this particular smell). From a Western point of view, there is nothing magical about the efficacy of plant baths; what we are presented with is only a very rational and astute manipulation and luring of the game animals' senses. But at the same time, Shipibo hunters maintain that a benevolent spirit is present in the smell and that it is this spirit that brings about good fortune in hunting activities. This kind of ethnographic finding challenges Malinowski's clear-cut dichotomy: when Shipibo hunters are playing with smells to attract good fortune and lure the game animals, it is not clear at all where the profane (i.e., practical knowledge) ends and where the sacred (i.e., magical beliefs and practices) starts.

To be sure, magic and science resort to different tools and use these tools to deal with different parameters, but it remains that both, in the end, target the same tangible outcome (e.g., a successful hunt). Magic is a kind of technology whose main purpose is to achieve practical efficiency; religion, on the other hand, is concerned with moral and soteriological purposes pertaining not to the mundane world but to what Nietzsche accurately dubbed a "backworld" (*Hinterwelt*).² For instance, magic is typically being used to ensure a plentiful hunt or crop while religion is used to ensure salvation or cosmological liberation. Stanley Tambiah reminds us that for Malinowski, "[r]eligious action was not like magic a means to an end, it was an end in itself and it celebrated ultimate values, such as Providence and Immortality" (1990, p. 69).

This relates to another important difference between magic and religion; magic, like science, envisions the world in a very mechanistic way, whereas religion claims that autonomous volitional beings can interfere within the mechanistic concatenations of the natural world. A miracle is thus defined as God's sovereign decision to temporarily suspend the mechanistic laws of nature. Clearly, there is no such a thing as a miracle in magic. When a magician performs some magical rituals, the summoned supernatural beings and forces are not at liberty to choose to please the magician or to rebuff her inquiry. By contrast, when one prays to God, one's wishes may perfectly be fulfilled *or not*. As Tambiah pleasingly remarks, the world of religion supposes a "sovereign God" while that of magic is replete with "manipulable divine beings" (Tambiah, 1990, p. 21). The classical view of anthropology, thus, is that "[i]n its quintessential form [...] magic is ritual action that is held to be automatically effective" (Tambiah, 1990, p. 7). Frazer has certainly been the most radical champion of this view. In the *Golden Bough*, he discusses at length the striking proximity between magic and science, i.e., the shared *mechanistic* view of the world and the commonly assumed idea that instrumental actions are

²For an evolutionary proposal as to how religion—with its otherworldly orientation—emerged, see: Baumard, Hyafil, Morris, & Boyer, (2015).

automatically effective.³ Evans-Pritchard—albeit more moderately than Frazer—does concur with the idea that magical rituals work to a large extent mechanistically. For example, he states that “[i]n asking a [magical] medicine to act on his behalf a man does not beseech it to do so. He is not entreating it to grant a favor. He tells it what it is to do, just as he would tell a boy were he dispatching him on an errand” (Evans-Pritchard, 1976, p. 177).

This concise review of some of the most important anthropological theories of magic has made clear that magic as practiced in real life is nothing mystical or irrational; it is an instrumental knowledge (although usually not evidence-based) and a technology involving the manipulation of supernatural entities in order to achieve very practical outcomes.

Magic in Practice: Three Case Studies

As reviewed earlier, the psychological approach holds that magic is mainly defined by its counterintuitiveness. Strangely enough, it is difficult to find this feature in the data on magic that anthropologists collect in the field (in natural settings). Cases typically discussed by anthropologists, in which some magic is at play, do not involve any impossible event or impossible being; on the contrary, such cases are strikingly ordinary. In order to illustrate this point, we will now present three cases which exquisitely epitomize the kind of phenomena anthropologists are concerned with when they speak of magic. Different mechanisms (i.e., different types of complexity drops) can trigger magical explanations. Each example provided in this section illustrates a specific type of complexity drop.

Case 1: Succession of Misfortunes

Let us start with an excerpt from Jeanne Favret-Saada’s captivating monograph depicting witchcraft in contemporary French countryside:

In the Bocage [...] ordinary misfortunes are accepted as “one-off”; so, a single illness, the loss of one animal, one bankruptcy, even one death, do not call for more than a single comment: “*the trouble with him is that he drinks too much*”; “*she had cancer of the kidneys*”; “*my cow was very old*”.

An onslaught by witchcraft, on the other hand, gives a pattern to misfortunes which are repeated and range over the persons and the belongings of a bewitched couple: in succession, a heifer dies, the wife has a miscarriage, the child is covered in spots, the car runs

³“[In both magic and science,] the succession of events is assumed to be perfectly regular and certain, being determined by immutable laws, the operation of which can be foreseen and calculated precisely; the elements of caprice, of chance, and of accident are banished from the course of nature.” (Frazer, 1922, p. 49). See also Ruth Benedict (1933, p. 40).

into a ditch, the butter won't churn, the bread won't rise, the geese bolt, or the daughter they want to marry off goes into a decline ... (Favret-Saada, 1980, p. 6)

The characteristics of magic identified by Favret-Saada are rather ordinary. People surmise that some witchcraft is involved not because they face an extraordinary event—e.g., an event violating some physical law—but rather because they detect an abnormal pattern in the way very ordinary events unfold. Occasional misfortune is not enough to suggest that some magical process is at work; persistent misfortune is required to infer that some witchcraft is involved. Magic characterizes the specific concatenations of events rather than the very content of events. In the case reported by Favret-Saada, the content of the events mentioned is in fact remarkably mundane and plainly natural.

Case 2: Unfortunate Relatives and Friends

The second case study is a classical and well-known anecdote provided by Evans-Pritchard in his monograph on Zande⁴ witchcraft:

In Zandeland sometimes an old granary collapses. There is nothing remarkable in this. Every Zande knows that termites eat the supports in course of time and that even the hardest woods decay after years of service. [...] it may happen that there are people sitting beneath the granary when it collapses and they are injured [...]. Now why should these particular people have been sitting under this particular granary at the particular moment when it collapsed? [...]. Through years it might have collapsed, so why should it fall just when certain people sought its kindly shelter? (Evans-Pritchard, 1976, p. 22)

In this passage, Evans-Pritchard pinpoints in a very concise manner several central features of witchcraft. First, magic unfolds in the world in a very ordinary way. The collapse of the granary does not violate any physical law. It is not an extraordinary event: real people believe in witchcraft not because they hallucinate witches flying on broomsticks but because they face in their life very ordinary facts which possess a specific structure (we will see below how this structure can be formally defined). Second, people who believe in magic do not live in a parallel outlandish and mystical world. They share with us exactly the same world and are very knowledgeable about the physical and biological processes governing the everyday world: they know that granaries usually collapse not because of supernatural powers but because of very mundane termites. What Evans-Pritchard finally shows us is that Azande people find it quite unbelievable that an unfortunate event—such as the collapse of the granary—could strike these very specific people at this very specific moment. Why them? Why then? If this unfortunate event harmed these very specific people, then, there must be a reason. In witchcraft logic, the reason is quite obvious: some ill-intentioned agent wanted to harm these very specific people.

⁴The Azande are an ethnic group of North Central Africa. In the 1920s and 1930s, Evans-Pritchard extensively studied Zande communities located in today's South Sudan.

The way the witch is thought to have interfered in the natural course of the world is certainly extraordinary (outstanding powers are required). However, Evans-Pritchard makes it clear that the specific mechanisms through which such witchcraft assaults are carried out remain largely opaque and Azande do not speculate about them. In fact, in their eyes, it is not of great importance to know how the witch did what he or she did. What matters is to be able to see this mundane event as the sign of some magical interference in the normal course of the world. And what is so remarkable is that this sign does not take the form of something extraordinary; it does not consist in the violation of physical laws. Only those who have learned how to read the ordinary unfolding of mundane events will be able to know that what happened has been partly brought about by some magical agency. Novices, on the other hand, will not even notice that this mundane event was magical in essence.

Case 3: Atypical Features Violating Factual Knowledge

The third case study differs from the two previous ones in several respects. To begin with, it concerns shamanism (in the Northern Asian area) rather than witchcraft. Moreover, this case illustrates how *beings* are identified as magical, as opposed to how *events* are identified as being such:

Tuvan hunters claim that there exist certain ‘special’ (*tuskai*) beasts in the taiga; these being rare, they are ‘sacred beasts’ (*ydyktyg aynar*) whom it is forbidden to kill. They can be recognized by distinct visible traits of very variable types: this could be a variegated deer or a white one (albinism); a stag with no antlers, or conversely with very large antlers; a doe with antlers; a bear with patches; or a white squirrel, sable or polecat; or a fish with just one eye. [...].

This particular treatment of singular animals is extremely widespread among the indigenous peoples of Northern Asia. [...]. ‘The Evenks and the Yakuts believe there to be in each animal species certain powerful specimens that have the qualities of a demi-god’ (Nikolaev 1961: 49). In this way, according to Yakut tradition, there exists among bears a ‘shaman bear’, which is marked out by its intelligence, its invulnerability, and by the patches on its fur (Kulakovskii 1979: 55). [...]. With ants there would also be a very large ‘mother ant’, while arctic foxes and hares also include shamans. According to the Nenets, the squirrel that keeps its red fur in winter is a shaman, and it is forbidden to kill it (Zelenin 1936: 208). (Stépanoff, 2015, pp. 172–173)

It seems that the cognitive mechanisms underpinning the detection of magical properties in this case are different from the ones at play in the two previous cases. In this passage, Charles Stépanoff points out that the ascription of magical powers to an entity does not always stem from the pattern of occurrence of a sequence of events (as in Case 1), or from the nature of the persons on which magical powers are exerted (as in Case 2), but it sometimes depends on the violation of some typical features (Case 3). In the examples discussed by Stépanoff, an external feature (color, size, shape, etc.) happens to be violating some ethological knowledge, and,

as a result, the bearer of the atypical feature is inferred to be endowed with some magical powers (as will be shown later, atypical features include not only atypical morphology but also atypical behaviors, smells, and sounds).

Magic and the Varieties of Counterintuitions

The relationship between magic and the violation of domain-specific and domain-general knowledge deserves to be further detailed. To begin, let us define a few key concepts. Cognitive scientists and psychologists often distinguish between domain-specific and domain-general knowledge. Domain-specific knowledge or intuitive ontology refers to a kind of intuitive knowledge concerned with a defined category of objects or processes. By “intuitive” we mean that this knowledge develops independently of any deliberate reflective effort (intuitive knowledge can thus be contrasted with reflective or scientific knowledge). By a “specific category” or an “ontology”, we mean that this knowledge does not apply to any thing. Importantly, intuitive domains, categories, or ontologies are parsed accordingly to our everyday use of cognition (and plausibly evolution) rather than by clear-cut folk or scientific categories: “the distinct cognitive domains [...] do not always correspond to real ontological categories [...]. For instance, the human mind does not draw the line between living and non-living things, or between agents and objects, in the same way as a scientist or a philosopher would do” (Boyer & Barrett, 2005, p. 98). An example of domain-specific knowledge would be the knowledge that living beings need to be fed in order to grow. This knowledge emerges early in childhood [when exactly it emerges is still a matter of debate (Atran & Medin, 2008; Carey, 1985)]; and this knowledge is “specific” to the extent that it strictly concerns biological objects and processes (as opposed to physical or psychological ones). In contrast, probabilistic knowledge applies to various kinds of objects or processes (Xu, 2007). For instance, it can be used both to infer the trajectory of a ball and to ascribe mental states to other people. It is hence domain general.

The definition of two additional concepts—core knowledge and modularity—is in order. Core knowledge refers to a set of cognitive mechanisms enabling young children (and sometimes even infants) to compute objects, persons, places, and numerosities in a quite sophisticated way. Each core knowledge system is domain-specific: it is dedicated to one specific category of objects and processes. Moreover, core knowledge is assumed to have a long evolutionary history (on core knowledge, see: Spelke, 2000). Modularity has been defined in several ways (Barrett & Kurzban, 2006; Fodor, 1983), but—to somewhat simplify the matter—it is a quasi-synonym of core knowledge: it refers to mechanisms enabling children to be innately endowed with some specific knowledge, or at least, to learn specific knowledge extremely quickly. Modules, it is argued, have been developed throughout evolution because they were, in some ancestral environment, adaptively

advantageous. As we can see, core knowledge systems and modules share many commonalities. The main difference between the theory of core knowledge (Spelke, 2000) and the theories of massive modularity (Barkow, Cosmides, & Tooby, 1992; Sperber, 1996) is that while the former considers that only about half-a-dozen core knowledge systems are to be identified, the latter claim that dozens of modules are to be recognized. Yet, in contrast to rational constructivism (Xu, 2007), both core knowledge and massive modularity agree that cognition is to a fair extent innate as well as domain-specific.

These clarifications can help us better understand the debate about the nature of magic. The three case studies just described suggest that magic does not tap into domain-specific mechanisms. This claim may seem disputable, though. Indeed, as we all know, magical explanations typically presuppose that mind can have an effect over matter, that diseases can be healed by thoughts, that vitality can be lowered by a spell, etc. Such explanations, it seems, blatantly violate domain-specific knowledge. Importantly, when we say that magic as studied by anthropologists is not based on the violation of domain-specific knowledge, we do not mean to deny that magical explanations often go against our domain-specific intuitions. The question we are addressing is different: we are interested in studying the *features* which lead people to interpret an event or a being as magical. Studying the *content* of magical explanations—rather than their *triggering conditions*—is vain as such explanations considerably vary within- and between-subjects. People recognize the magical character of events in a remarkably consistent way; by contrast, they do not invest much effort in speculating what are specific and detailed magical processes which brought about such events—or, when they do (for example, to please the ethnographer), they are remarkably inconsistent through time. In the remainder of this chapter, we will thus restrict our investigation to the cognitive mechanisms through which an event or a being is detected as magical—nothing more.

The common claim of developmental psychology is that features triggering magical explanations can be characteristically defined as violating some domain-specific knowledge or some intuitive ontology. It is thus claimed that people resort to magic when faced with events or entities which contradict intuitive expectations as to how physical, biological, or psychological beings should behave. This is clearly stated by the editors of the state-of-the-art book *Imagining the impossible*: Developmental psychologists, they notice, are mainly concerned with the study of thinking about violations of domain-specific constraints and thinking going beyond the ordinary (Rosengren, Johnson, & Harris, 2000, p. xiii). This type of magic contrasts with the type of magic that anthropologists have been documenting at length in their monographs. Strikingly enough, in the three case studies that we have examined, no obvious violation of physical, biological, or psychological laws is observed. In Case 1, magic stems from the occurrence of very mundane physical and biological events (e.g., a car accident or the death of an animal). Similarly, in Case 2, the event reported to be magical is the collapse of a granary. Clearly, as such, this collapse does not violate any domain-specific knowledge. Case 3 does feature a violation, but what is being violated is

domain-general knowledge as opposed to domain-specific knowledge.⁵ In the end, none of the three examined cases contradict our intuitive ontologies. It must be consequently acknowledged that in real life, people resort to magical explanations in front of mundane events, and not impossible ones.

It is worth elaborating a bit more on Case 3. Indeed, influential cognitive theories have advanced that entities and properties violating knowledge more or less massively are constitutive of supernatural thinking (Pyysiäinen, 2002). Therefore, it remains to be determined whether cases such as Case 3 can be best characterized as a *minimal* violation of domain-general knowledge or a *maximal* violation of knowledge. For the sake of clarity, let us define four types of counterintuitions:

- *Type 1. Minimal domain-specific counterintuitions* are triggered by objects or events which possess one or very few properties violating domain-specific knowledge (Barrett, 2004b; Boyer, 2001). A human able to fly or a mountain able to talk are good illustrations of what minimal domain-specific counterintuition amounts to.⁶
- *Type 2. Multiple domain-specific counterintuitions* are triggered by objects or events which possess many properties violating domain-specific knowledge. Here is an example: “a dog that was made in a factory gives birth to chicken, can talk to people, is invisible, can read minds, can walk through walls, and can never die” (Barrett, 2004b, p. 23). Unlike maximal violations (see Type 4), multiple domain-specific violations are not intrinsically contradictory or utterly abstruse.
- *Type 3. Minimal domain-general counterintuitions (or oddities)* are triggered by objects or events which possess one or very few properties violating domain-general knowledge (Gille, 2014). For example, a cat with one eye or a pink cat (as opposed to, say, a flying cat) minimally violates *domain-general* knowledge but does not violate *domain-specific* knowledge. Let us spell this out: a cat qua physical object is domain-specifically known to be subject to gravitation; a cat qua living being is also domain-specifically known to grow and die. Any violation of this kind of specific and theory-based knowledge (e.g., a flying cat or an immortal cat) falls under Type 1 described above. Now, in addition to this specific theory-based knowledge, people have domain-general prototype-based knowledge. For example, cats have typically two eyes and four legs. Importantly, this knowledge cannot be inferred from knowledge that cats are living beings or physical objects. It is learned through domain-general mechanisms rather than inferred from domain-specific knowledge. When this

⁵On the distinction between domain-general (or factual, or prototypical) knowledge and domain-specific (or ontological) knowledge, see: Barrett (2004a, b), Chap. 2, Gille (2014).

⁶Minimally counterintuitive concepts are created as follows: “First, take an ordinary concept, such as “tree”, “shoe”, or “dog”, that meets all of the naturally occurring assumptions of our categorizers and describers. Then *violate* one of the assumptions. For instance, as a bounded physical object, a tree activates the nonreflective beliefs governing physical objects, including being visible. So make the tree invisible [...], and you have an MCI [minimally counterintuitive concept]” (Barrett, 2004b, p. 22).

prototypical domain-general knowledge is violated (e.g., a cat with one eye or two legs), this results in Type 3 counterintuition.

- *Type 4. Maximal domain-general and/or domain-specific counterintuitions (or contradictions)* are triggered by objects or events which are intrinsically contradictory and utterly abstruse and/or whose properties massively violate domain-specific and/or domain-general knowledge (Atran, 2002; Sperber, 1975). A good example of a contradictory concept is the Christian God, who is purported to be both trine and one.

Clearly, Type 1 and 2 do not apply to any of the three cases under examination. What about type 4? Now that we have clearly contrasted Type 3 counterintuition with Type 4 counterintuition, it seems obvious that only the former accurately describes what is going on in Stépanoff's reports. Bears and squirrels purported to possess magical powers are not ascribed these powers because they are able to fly in the sky (minimal domain-specific counterintuitiveness); neither are they ascribed such powers because they are massively contradictory to the point that it is hard to conceptualize them (contradiction). What makes these animals magical is simply that their features include some unusual color or some congenital malformation (minimal domain-general counterintuitiveness or oddities).

In cases described by ethnographers, events or beings typically recognized as magic are either not counterintuitive whatsoever (think of the collapse of a granary, in Case 2) or are at most minimally and domain-generally counterintuitive (think of the albino squirrel, in Case 3). It is often claimed that the domain-specific or modular architecture of the mind is key to understand magic. In contrast to such claims, we are left with the suggestion that the central component of magic lies in the domain-general interpretation of ordinary events. Importantly, the modular view of the mind is of no avail when it comes to understanding such domain-general cognitive processes. On the other hand, probabilistic models of the mind (Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Tenenbaum, Kemp, Griffiths, & Goodman, 2011; Xu, 2007) seem particularly well suited to do the job. This is not to say that the modularist view of the mind should be discarded and will never be able to help us understand anything about magic. The matter is more intricate. Indeed, as shown in the beginning of the chapter, psychological experiments demonstrate that people do explain in magical terms domain-specific counterintuitive events; that is, when presented with such events, almost all children and at least some adults do resort to magical explanations. On the other hand, when people are not presented with laboratory stimuli but are freely navigating in the real world, what they typically explain in magical terms are domain-general minimally counterintuitive events as well as perfectly intuitive events. Domain-general models of cognition, it seems, are thus better equipped to understand when and how magic is evoked. The apparent contradiction between the two strands of evidence just discussed will be explored and adjudicated later in the chapter. For now, some formalization work is required in order to better understand the meaning and significance of the anthropological data under scrutiny.

Formalizing Anthropological Data: The Probabilistic Model of Magic

Complexity Drop as a Cue to Agency

Anthropological evidence demonstrates that people resort to supernatural explanations in front of events or beings which do not violate any domain-specific intuition. Models of magic based on domain-specific mechanisms are thus of no avail to understand such anthropological data. What is therefore needed is a model of magic which explains how magical thinking is underlain by domain-general cognitive mechanisms. As we will see, Kolmogorov's concept of algorithmic⁷ complexity, or rather the revamped version of this concept (Chater & Vitanyi, 2003; Dessalles, 2013; Griffiths & Tenenbaum, 2003; Vitanyi & Li, 2000), will prove particularly useful. What follows is a brief summary of a general model of magic that one of us has developed at length elsewhere (for more details, see: Fortier, in preparation).

Kolmogorov's theory of complexity aims at understanding the nature of randomness. The proposal is that randomness amounts to maximal complexity. Complexity, in turn, is defined by size of the shortest program that generates a given string; the longer the program (the higher the number of instructions a program needs to generate a given string), the more complex the program. For example, more instructions will be needed by a program to generate this string: 4, 7, 3, 5; as compared to that string: 1, 2, 3, 4. A random string such as (4, 7, 3, 5) demands more instructions because it is not compressible. In Kolmogorov's terms, a compressible string is less complex than a noncompressible one.⁸

Admittedly, one apparent limit of Kolmogorov's theory of complexity is that it is not computable. However, as Jean-Louis Dessalles notices, as long as we are using complexity in the field of cognitive science, this theoretical limit is harmless since we are by definition dealing with specific and limited "computing machines" (i.e., human minds) (Dessalles, 2007). When applied to human cognitive systems, acknowledging that complexity is a relative concept is not a problem anymore. Concretely, this means that what is complex for a given cognitive system (for person A) will not necessarily be complex for another system (for person B). For instance, (1, 9, 7, 4) may look totally random (highly complex) to person A, but it may simultaneously look very simple (not random at all) to person B, because, say, person B was born in 1974. Similarly, the first ten digits of π may look totally random (highly complex) to a non-mathematician and yet they look very simple (not random at all) to a well-trained mathematician.

⁷Henceforth the concept of "algorithmicity" (and its adjective "algorithmic") will be only used in a Kolmogorovian sense. The term strictly refers to the algorithmic information theory and the calculation of the complexity of objects (such as strings).

⁸For details on the concept of compressibility, see: Li and Vitanyi (1997, p. 108 et sq.)

Complexity theory has been fruitfully applied to various domains of cognitive science. In particular, Dessalles has developed a wide-ranging theory of unexpectedness and relevance entirely based on Kolmogorov complexity (Dessalles, 2007, 2010, 2013; Saillenfest & Dessalles, 2015). The tenet of his theory is that unexpectedness can be defined as the ratio between expected complexity and observed complexity. More specifically, if an event which was expected to be complex turns out to be very easy to describe, it will be deemed highly unexpected. If, for instance, I am traveling in Patagonia, I expect people I meet during my trip to be complex people (i.e., to be strangers whose description cannot be simplified or compressed); now, if I happen to meet my neighbor or a university colleague or a celebrity (i.e., a person whose description can be easily compressed and whose description is therefore simple), I will certainly be baffled by the unexpectedness of the encounter: I was expecting to encounter someone complex to describe and I happened to encounter someone unexpectedly simple. Unexpectedness can thus be defined as follows:

$$U = C_{\text{exp}} > C_{\text{obs}}$$

(where U stands for unexpectedness; C_{exp} for expected complexity; and C_{obs} for observed complexity).

This simple but powerful equation interestingly echoes a recent line of research in developmental psychology exploring the relationship between nonrandomness (or simplicity) and agency (Keil & Newman, 2015; Kushnir, Xu, & Wellman, 2010; Ma & Xu, 2013; Newman, Keil, Kuhlmeier, & Wynn, 2010; Wellman, Kushnir, Xu, & Brink, 2016). These studies show that very early on, humans explain the absence of randomness (i.e., high simplicity) by postulating the interference of an agency. This inference is particularly strong, as 9-month-olds (Ma & Xu, 2013) seem already able to hypothesize that nonrandom outcomes are generated by human agents and random outcomes by mechanistic processes. It appears that abnormal order (violation of randomness) is heuristically used as a cue to agency. Such a relationship seems indeed quite intuitive. If I am walking in a meadow, in a remote location, and I see flowers randomly scattered around me, there is nothing surprising about it. But let us now imagine that I suddenly encounter a perfectly straight line of flowers right in the middle of a remote meadow. There is no doubt that I will be surprised and think: someone must have planted these flowers here this way, this cannot be the result of randomness alone! So, in such a case, I was expecting high complexity (disorder) and I observed low complexity (order); as a result, I inferred that an agent must have somehow interfered: an agent must be added to the picture in order to explain the complexity drop. Hence the idea that complexity drop functions as a cue to agency:

$$(C_{\text{exp}} > C_{\text{obs}}) \rightarrow AD$$

(where AD stands for agency detection).

A further nicety has to be added to this model of agency detection. In an ingenious study, Falk and Clifford Konold (1997) have shown that the difficulty to encode a sequence (i.e., the complexity of the sequence) is a very good predictor of judgments of apparent randomness. In other words, probabilistic judgments—judging whether a sequence was generated by a random or an intentional process—are based on the experienced ease (or difficulty) in encoding a sequence. This explains why people are sometimes biased in their probabilistic judgments: the most difficult sequence to encode is not always the most random one. Although subjective complexity is not always a good guide to objective randomness, it is heuristically used by subjects to guess what the nature of the underlying generating process is.⁹

To summarize, the merit of Dessalles's theory is to provide a rigorous computational framework making straightforward predictions as to which events (or entities) will be deemed unexpected. Moreover, the developmental literature just reviewed helps us understand how unexpected simplicity is used as a cue to agency from an early age. Finally, Falk and Konold's work, as well as studies on metacognitive feelings, specify the link between experienced ease (i.e., fluency) on the one hand, and simplicity and nonrandomness on the other; so doing, they enable us to grasp the detailed psychological mechanisms through which Dessalles's computational framework is likely to be implemented. The following thesis can now be put forth: *when an event (or an entity) is expected to be complex (i.e., to be disfluent or difficult to process) but when it turns out to be simple (i.e., to be fluent or easy to process), this event (or this entity) will be perceived as unexpected, and, if the observed complexity drop (the gap between expected disfluency and experienced fluency) is big enough, it will be inferred that an agent must have caused the observed reduction of complexity.*

The Complexity Drop Model of the Supernatural (CDMS)

The previous section explains how *human agency* is detected. However, understanding magic requires us to explain how *supernatural agency* is usually detected. In order to grasp the mechanisms of supernatural agency detection, let us go back to Case 1. As previously explained, Favret-Saada's book demonstrates that, in the Bocage, the magical nature of an event lies in the very structure of a sequence of events rather than in the intrinsic features of each single event. If a misfortune happens to someone (e.g., a heifer's death), it will be interpreted as plainly natural.

⁹It is worth noting that the experienced ease (or difficulty) in encoding Falk and Konold are referring to is closely related to (if not synonymous with) the *feeling of fluency* (Unkelback & Greifeneder, 2013). This concept, which has extensively been discussed in the field of metacognition, is usually defined as the ease of information processing. Feelings of fluency are notably characterized by their phenomenology (Reber, Fazendeiro, & Winkielman, 2002; Schwarz & Clore, 2007).

However, if this misfortune is preceded and followed by other unfortunate events, each of these events will then be understood not only as the result of a natural process (e.g., the heifer died because of illness) but also as the obvious manifestation of some bewitchment.

We can now easily make sense of Favret-Saada's case in the light of the model adumbrated above. People generally expect events, either fortunate or unfortunate, to occur randomly. Let us write "1" for any fortunate event and "0" for any unfortunate event. People's expectation, then, is that any sequence of their life will look more or less like this: 010010110100. But sometimes, as we all know, life looks more like that: 000000000000. This is precisely what happens in Case 1. In our own terms, bewitched peasants are reasonably expecting a fairly complex sequence of events to take place (010010110100), but what effectively takes place happens to be an abnormally simple pattern of events (000000000000). The detection of some supernatural agency usually coincides with a striking complexity drop. However, the precise reason why complexity drops lead to the detection of *supernatural* agency remains to be spelled out.

To understand this, let us suppose that we are playing with a die. We expect the outcome of throws of a die to be random (for example: 3, 6, 1, 4, 2). Now, let us imagine that the observed outcome happens to be the following: 6, 6, 6, 6, 6. According to our model, this leads to some agency detection: someone must have tampered with the die—the die is definitely not a fair one! In front of such a complexity drop, the abnormal simplicity of the sequence is explained by resorting to an agent, and importantly this agent is a human one (not a supernatural one) because we possess a *natural causal schema* specifying how humans can tamper with dice. Let us now return to Case 1. As with the loaded die, a complexity drop is observed (a random sequence was expected and a nonrandom one occurs). Unlike the case of the loaded die, however, in Favret-Saada's case, we do not have any *natural causal schema* at our disposal in order to explain how a human agent caused the complexity drop. We know very well that humans can tamper with a die whereas we know of no human-like being able to control the course of existence and able to turn fortune into misfortune (or the other way around). When facing a complexity drop such as the one discussed by Favret-Saada, we are thus inclined to explain the nonrandom sequence of misfortunes by postulating the existence of an agent, and because no natural agent can turn fortune into misfortune, this agent must be a supernatural one. Indeed, only a *supernatural causal schema* can do the job of explaining such a complexity drop. Two kinds of complexity drop should thus be distinguished: one which can be accounted for by *natural causal schemas* and another which can only be accounted for by *supernatural causal schemas*. Hence, the two criteria of the Complexity Drop Model of the Supernatural (CDMS): *an event or a being will be ascribed supernatural features, if (i) the complexity which defines it is much lower than what was expected, and if (ii) the observed complexity drop cannot be accounted for by some natural agency.*

The CDMS seems able to explain fairly well the supernatural agency detection at work in Case 1. But how well can it equally easily account for cases such as those discussed by Evans-Pritchard (Case 2) or Stépanoff (Case 3)? It is here that

Dessalles's theory proves particularly fruitful and powerful. Indeed, not only does Dessalles provide a mathematical framework defining complexity drop in a thoroughgoing manner, but he also identifies distinct types of complexity drop.¹⁰ As will be shown, each of the three case studies presented above illustrates one specific type of complexity drop. For the sake of clarity, let us define the three types of complexity drop:

- *Type 1. Complexity drop triggered by redundancy.* First, a complexity drop can stem from the simplicity (the redundancy) of a pattern of events: one is expecting a random (a complex) string (e.g., a string where 1s and 0s are instantiated randomly) but observes a redundant (a simple) string (e.g., a string consisting only of 1s or 0s).
- *Type 2. Complexity drop triggered by proximity/familiarity.* Second, complexity drop can stem from the abnormal simplicity of the place or the characters involved in an event. Famous places (e.g., the White House) or famous people (e.g., Barack Obama) are remarkably simple, because computational resources required to describe them are very limited. Similarly, familiar places (the place where one was born) or familiar people (one's relatives) are particularly simple (not for everyone, but for the person concerned).
- *Type 3. Complexity drop triggered by atypicality.* Third, complexity drop can stem from atypical features. Let us consider the following geometric shapes: 3 squares and 1 chiliagon. To be sure, a chiliagon is intrinsically much more complex than a square (describing it requires more computational resources). However, if we want to single one specific shape out of the four aforementioned, it will be simpler to single out the chiliagon than any of the three squares. Singling out the chiliagon requires the program to follow one single instruction—"pick up the shape which does not look like the others"—whereas singling out one of the squares requires our program to use an additional instruction—"pick up the shape which looks like the others" and "which is located on the left/right." This is why, from an algorithmic standpoint, atypical items are intrinsically more complex than typical ones but are paradoxically simpler to describe.

As demonstrated earlier, Favret-Saada's case illustrates how Type 1 complexity drop can lead to the detection of supernatural agency. It is worth emphasizing that Type 1 complexity drop can result from either an abnormal succession of unfortunate events (Favret-Saada, 1980) or an abnormal succession of fortunate events (such as being abnormally successful at hunting or in love affairs) (Brown, 1986). We now turn to the two other cases.

Evans-Pritchard insists that when Azande resort to magical explanations, they do not mean to deny the reality of physical or biological processes at work (e.g., the

¹⁰See: Dessalles (2007, 2010). As a matter of fact, Dessalles identifies about a dozen of types of complexity drops. In this chapter, we discuss only three of them. However, several types of complexity drops singled out by Dessalles share commonalities and could therefore arguably be grouped together.

termites eating the wood of the granary) but they want to explain why is it that these “particular people”—these close friends and these relatives—have been injured and killed (Evans-Pritchard, 1976, p. 22). Clearly, what puzzles the Azande is not that a granary can fall—they know that, in virtue of the laws of nature, such a thing can very well occur—but rather that the granary can fall at the very wrong time and on the very wrong people. This is too big a coincidence! According to the CDMS, there is nothing incomprehensible in the Azande’s puzzlement. The event they interpreted in magical terms is indeed puzzling because the people concerned were very “particular people” (relatives and close friends); these people were, in algorithmic terms, abnormally simple people. So some (supernatural) agency had to be held accountable for this complexity drop. Evans-Pritchard’s case thus illustrates how a Type 2 complexity drop can lead to the detection of supernatural agency.

In Stépanoff’s case, we are dealing with entities (mostly animals, but also plants and rocks) whose features are remarkably atypical (because of their color, size or shape). This schema fits very well with what we have defined as Type 3 complexity drop. Atypical entities (such as an albino squirrel) are algorithmically simpler to describe than entities sharing many features with the other surrounding entities. It is worthy to note that Type 3 complexity drop can be brought about by an atypical *physiological* feature but also by an atypical *behavioral* feature. Anthropologists have extensively documented how, in certain regions of the world, animals behaving abnormally are suspected to be not real animals but spirits momentarily inhabiting an animal or taking the deceiving shape of an animal (e.g., Keifenheim, 1999).

Let us take stock: when an event or a being is processed with fluency (when it is algorithmically simple) while it was expected to be processed with difficulty (expected to be algorithmically complex), the unexpected fluency (the complexity drop) is explained by inferring that some agency has interfered at some point in the generating process; when, furthermore, the kind of agency required to cause such an unexpected fluency (such a complexity drop) does not fit with ordinary causal schemas, an extraordinary (or supernatural) causal schema is then posited or used. It must also be remembered that unexpected fluency (complexity drop) can be caused by distinct mechanisms (for a summary, see: Table 2).

Comparing the CDMS with Other Cognitive Models of Magic

In recent years, several models of magic have been developed by cognitive scientists of religion.¹¹ Spelling out how our proposal distinguishes itself from other well-established models will enable us to better understand the originality and scope of the CDMS.

¹¹Note that cognitive science of religion is not restricted to the study of religion *stricto sensu*. Many researchers working in this field are investigating shamanism, magic, witchcraft, etc.

Epidemiological Versus Doxastic Models of Magic

Among the most influential models of magic, those accounting for the epidemiology of supernatural representations deserve a special mention (Atran, 2002; Barrett, 2004b; Boyer, 2001; Sperber, 1996). Even though they differ in the details, all these models share the core thesis that owing to their counterintuitiveness, supernatural representations prove particularly catchy and can thus spread very successfully across space and time. The cognitive science of religion thus seems to corroborate the psychological approach to magic: recognizing that magical representations are essentially counterintuitive is apparently quite consistent with saying that counterintuitive stimuli tend to trigger magical explanations. Nevertheless, there is good reason to resist conflating these two lines of research and to maintain that the counterintuitive models of magic developed by the proponents of the epidemiological approach do not accord well with those put forward by developmental psychologists.

What psychologists demonstrate is that children (and, in some cases, adults) are inclined to postulate the existence of magical processes or magical beings in order to explain the occurrence of impossible events. That is, the need for an explanation leads people to be *committed to the existence* of magical entities and to *believe* in magical entities. In this regard, psychological models of magic attempt to explain the *doxastic* dimension of magic. By contrast, epidemiological models of magic are impervious to doxastic and ontological issues. Their aim is not to explain why or even whether people believe in or are ontologically committed to certain entities. Epidemiological models modestly aim at explaining why certain representations are very memorable and easily transmitted and why others are less so. Now, *pace* Boyer, remembering is not believing.¹² For example, we may all very well remember the concept of Santa Claus; the concept of Santa Claus may be very catchy; still, most of us do not believe in the existence of Santa Claus. The folklore is replete with myths, legends, and tales depicting counterintuitive entities. Fair enough. But this does not mean that representing (remembering, narrating, painting, etc.) these entities amounts to being committed to their existence—some people believe in some of these entities and others do not. Epidemiological models attempt to understand how representations spread in space and time but they have nothing to tell us as to why people might believe (or not believe) in them. Epidemiological models remain silent about the doxastic dimension of magic; in this regard, they are very different from the models of magic that developmental psychologists advocate.

Epidemiological models tend to equate supernaturalness with counterintuitiveness. Yet, as just mentioned, people can very well represent counterintuitive entities and still not believe in them. The set of counterintuitive representations is therefore bigger than that of supernatural beliefs. Moreover, as demonstrated earlier, people often believe in magic in situations in which no counterintuitive event or entity is

¹²Cf. Boyer (2001, pp. 302–304). For a critical discussion of Boyer's assumptions, see Fortier (in press).

involved (think of Favret-Saada's or Evans-Pritchard's case). It is important to understand why people interpret certain events (e.g., a sequence of misfortunes) or certain beings (e.g., an albino squirrel) as magical and why they are committed to the reality of these representations. Remarkably, the CDMS meets all these challenges.

The Hypersensitive Agency Detection Device and the CDMS

Some proponents of the epidemiological account of magic have attempted to meet the doxastic challenge by developing a model which explains why people are sometimes inclined to be committed to the existence supernatural entities. This is notably the case of Justin Barrett, who endorses the epidemiological program, but who has also developed the Hypersensitive Agency Detection Device (HADD) model in order to explain why people *believe* in some of the representations that they entertain (Barrett, 2004b). The HADD model is based on a wealth of studies showing that humans (and, to a certain extent, primates) are very adept at detecting agency (see: Rutherford & Kuhlmeier, 2013). By and large, the HADD assumes that it is evolutionarily advantageous to overdetect danger and agency in the world: it is better to trigger false positives (to detect a predator when there is none) rather than omitting actual signals (missing to detect a predator which is there). According to this view, humans have acquired a special ability to detect agency in their environment, and as a result, they are now prone to see agents everywhere (to see faces in the clouds, to attribute intentions to beings devoid of mental states, etc.). The HADD model is wide ranging. It includes three distinct versions of the same general claim:

- *From misperception to agency detection.* This line of research has been mainly explored by Stewart Guthrie (1993). Pareidolia is a striking instance of the kind of agency detection. Humans are prone to see anthropomorphic patterns (typically, a face) where there is no such pattern. Arcimboldo's paintings exemplify this bias fairly well. If one sees a face in a rock, one may subsequently infer that this rock is a living being and possess some kind of agency.
- *From motion to agency detection.* Another type of agency detection is based on motion. Heider and Simmel's (1944) famous study provides a good illustration of this phenomenon. When one sees a plain shape (such as a triangle) moving in a certain way, one cannot help attributing mental states to this triangle (e.g., this triangle is attacking or protecting someone). Motion is used as a cue to agency; as a result, agency can be ascribed to objects which are known for not possessing any agency.
- *From etiological gaps to agency detection.* Barrett (2004b) has extensively discussed cases where an event occurs while no visible agent is present. A door closing by itself generates an etiological gap: it is known that doors don't close by themselves, if it seems to do so, it must be because of an invisible agent. In

such a case, the existence of an invisible agent will be postulated to elucidate why the door closed itself and to fill the etiological gap.

Can these three versions of the HADD account for the three case studies introduced earlier? Does the HADD differ from the CDMS? Let us first focus on the two first versions of the HADD (misperception and motion). The weakness of these two models is that at most they explain how *natural agency* is ascribed but they do not say a word as to how *supernatural agency* is ascribed. By contrast, the CDMS readily explains how, for example, supernatural agency is inferred from a sequence of unfortunate events (only an agent with supernatural powers is able to turn fortune into misfortune). How could supernatural agency be inferred from pareidolia or motion? Besides, these two versions of the HADD wrongly assume that agency detection can only occur if the agent is *visible*. Yet, in many cases, supernatural agency detection occurs while the agent is not visible whatsoever (in Bocage witchcraft, for instance, misfortunes are visible but the causes of these unfortunate events—the supernatural agents—are not).

The third version of the HADD is a little more promising. *Inter alia*, it manages to explain how detection of *supernatural agency* occurs.¹³ The main drawback of this model, however, lies in the fact it takes agency detection to follow from the violation of physical laws. In the example of the door, an etiological gap needs to be filled only because some domain-specific counterintuition occurs. This is a serious problem because, as explained earlier, most of the cases provided by ethnographers do not involve the violation of any domain-specific knowledge. The HADD fails to account for this crucial fact. The CDMS, on the other hand, perfectly meets this challenge.

Probabilistic Models of Magic and the CDMS

Finally, it has been proposed that probabilistic models could help us understand the main mechanisms of magic (Bronner, 2003, 2007; Clément, 2003). Building upon previous research on probabilistic reasoning, it has been suggested that a cognitive bias such as the *gambler's fallacy* could explain why people have a distorted perception of probabilities and randomness, and thereby, why they resort to magical explanations. As Bronner (2003, 2007) remarks, people wrongly assume that the usual random course of nature is not compatible with homogeneity¹⁴; if a

¹³It is worth remarking that the supernaturalness of the agency at hand is still quite limited: admittedly, no natural agent can invisibly close a door, but it remains that closing a door, unlike turning fortune into misfortune, is not an outstanding feat.

¹⁴This is Bronner's own term. It must be pointed out that what is meant exactly by "homogeneity" remains rather vague. The term clearly seems to overlap with the concept of simplicity. The problem, however, is that Bronner does not provide any formal mathematical definition of homogeneity. As a consequence, no straightforward prediction can be made. By contrast, Dessalles' theory and the CDMS are grounded on neat formalism.

succession of fortunes (or misfortunes) occurs in a row, people will be puzzled and will look for a special (a supernatural) explanation, because they will reason that such a succession cannot be brought about by nature alone. This probabilistic approach to magic has many ideas in common with the CDMS.

Despite their valuable strengths, however, Clément's and Bronner's models remain limited for at least two reasons. First, they account very well for people's tendency to misrepresent randomness and probabilistic processes in general, but they do not spell out the link between probabilistic reasoning and agency detection. The link between cognitive biases and supernatural agency detection remains unexplained. The merit of Kushnir, Wellman, Xu, and colleagues is precisely to have discovered what the missing link was. Drawing upon these studies, the CDMS clearly spells out how supernatural agency detection stems from probabilistic reasoning. This is an important difference with Clément's and Bronner's models.

The second limit of existing probabilistic models of magic is that not any improbable event is liable to be interpreted as magical. Admittedly, in the three case studies we have examined something highly improbable is happening (i.e., an uninterrupted sequence of misfortunes is improbable; having relatives killed by the collapse of a granary is improbable; meeting an atypical animal or plant or rock is improbable); yet, there are plenty of improbable events which will never be interpreted in magical terms. Improbability is a necessary but not a sufficient condition of magic. Unlike strictly probabilistic models of magic, the CDMS does not predict that any improbable event triggers magical explanations; more specifically, it states that only *complexity drops* are likely to trigger magical explanations. *Complexity surges* (i.e., encountering randomness when order was expected) are very often highly improbable but they are not interpreted in magical terms. Think of the case mentioned earlier: I am walking in a meadow in which flowers are scattered randomly, and when I face an abnormally straight line of flowers I infer that a human agent must have intentionally planted the flowers this way. Now, let us consider the reverse case: I am walking in a *jardin à la française* and I am expecting to encounter only straight lines of flowers around me, but then, suddenly, I encounter flowers randomly distributed. Would I infer from this complexity surge that some human agent has intentionally planted these flowers randomly? This is very doubtful. Hence, the idea that agency detection is specifically a matter of complexity drop and not a matter of statistical improbability. The CDMS is more parsimonious and specific than existing probabilistic models of magic and it is able to make predictions which better fit with actual anthropological data.

Toward a Big Picture: Putting Counterintuitive-Magic and Probabilistic-Magic Together

Secularization and Explanatory Coexistence: The Persistence of Magic Through History

Before examining how the types of magic identified so far are combined and employed through development and cultures, it will be useful to look at the historical trajectory of magic. In particular, did the advent of modernity and secularization take place to the detriment of magic? It is often claimed that modernity and magic are antagonistic. For example, Weber (2003) and Thomas (1971) have both championed the view that religion and science have considerably contributed to the elimination of magic. Thomas notably maintained that the advance and development of science has gradually confuted magical beliefs; similarly, Weber famously contended that the world is now disenchanted (*entzaubert*) and there is no room in the secular cosmos for magic anymore. The accuracy of such assertions obviously depends on which definition of magic is adopted. If magic is defined in counterintuitive terms, then there is perhaps a grain of truth in Weber and Thomas' claim. For instance, the ingenuousness that people showed in front of counterintuitive feats a few centuries ago (Le Bouyer de Fontenelle, 1687) has arguably largely vanished; it might be that now more than before, adults see impossible events as being the result of trickery rather than real magic (but see Subbotsky, 2004, 2014). If magic is defined in terms of complexity drops rather than impossible events, however, Weber and Thomas' thesis might be proved wrong. In the remainder of this section, we will show that magic (as defined by the CDMS) is still widespread in the modern world. We will also suggest that two types of explanatory coexistence should be distinguished: *forced coexistence* (when two competitive explanatory frameworks coexist) and *easy coexistence* (when two non-competitive and complementary explanatory frameworks coexist).

It must first be acknowledged that the so-called wane of magic has never properly taken place. Human activities in which a great deal of uncertainty is involved are still, even in the modern world, replete with magical beliefs and practices (Bersabé & Martínez Arias, 2000; Blumberg, 1963; Keinan, 1994; Rudski & Edwards, 2007; Vyse, 1997). A plethora of anecdotes demonstrate that politicians and sportsmen (two domains in which uncertainty is prominent) often resort to the service of magicians (Bleak & Frederick, 1998; Blumberg, 1963; Pégard, 2000). Moreover, secularization processes are often only illusory. For example, one rather convincing idea that Favret-Saada (1986) has put forward is that the rise of witchcraft observed in the Bocage was a direct consequence of the demagification of the Catholic Church. Following the demands of clerical authorities, the priests stopped indulging in magical rituals, and as a result, Bocage peasants felt the pressing need of dealing with their existential misfortunes and they thus turned to witchcraft. The secularization of religion (in this case of the Catholic Church) was

not a sign of the secularization of the world; it simply amounted to a transfer of magical beliefs and practices from one side (the Catholic Church) to another (witchcraft).

It has been shown above that magic is to be distinguished from religion. While the latter is mainly concerned with otherworldly matters and big questions, the former is remarkably pragmatic, utilitarian, and down-to-earth. It is a pity that recent research on explanatory coexistence—i.e., on the simultaneous use of natural and supernatural explanations—does not subdivide the broad category of the supernatural into to subcategories: magic and religion (Legare, Evans, Rosengren, & Harris, 2012; Shtulman & Lombrozo, 2016). In itself, explanatory coexistence challenges the secularization story; if people explain the world in both natural and supernatural terms, this implies that, after all, supernatural beliefs have not entirely disappeared. It is worth emphasizing, however, that the thesis of the disenchantment of the world concerns chiefly magic. Modernity and secularization are claimed to be antagonistic with magic but not necessarily with religion (Weber, 2003). For instance, the coexistence of evolutionary and creationist explanations in accounts of the origin of man does not directly contradict the secularization story (Evans, 2001), whereas the coexistence of biological and witchcraft explanations in accounts of illness clearly does (Legare & Gelman, 2008).

Only tangible efficiency and material well-being matter in magic; there is no room for morality or otherworldly speculations. Science and religion, on the other hand, do address speculative issues such as the origin of man, of the universe, or the fundamental nature of things. Science shares with magic its instrumental and mechanistic image of the world, but it shares with religion its taste for abstract theorizing and big questions. In science and religion, a great amount of reflective effort is put in refining theories and grasping truth. By contrast, magic does not exhibit any kind of craving for truth; theoretical refinement is not seen as intrinsically valuable and worth pursuing; magic invests efforts in theorizing only to the extent that it will bring tangible outcomes.

On issues such as the origin of the universe, the earth, and man, religion used to hold theories directly competing with those of science. For example, the Roman Catholic Church has been forced to gradually update its theories in response to numerous challenges sparked by science (Hess & Allen, 2008). By contrast, magic is seldom in competition with science. Scientific explanations are not meant to elucidate why, for example, a complexity drop takes place when a granary collapses (Case 2). Magical explanations, on the other hand, are chiefly concerned with complexity drops. Magical theories are unlikely to be defeated by science because science does not intend to build any competitive theory accounting for complexity drops.¹⁵ The coexistence of magical and scientific explanations seems to be very

¹⁵It could be objected that sometimes science does produce theories liable to be competitive with magic and to demystify it. For example, when Gilovich, Vallone, and Tversky (1985) attempt to demonstrate that the “hot hand” phenomenon is a sheer fallacy, it seems that their scope indeed overlaps that of magic. But two remarks are in order: first, it is important to underline that scientific investigations demystifying magic are much rarer than those demystifying religion; second,

natural: each has a specific and an exclusive explanatory scope. No strenuous theoretical work is required to make them compatible. By contrast, the coexistence of religious and scientific explanations requires a great amount of theoretical refinement and hybridization. Making science and religion compatible is hard work because their explanations are competing.

The two distinct patterns of coexistence that we are trying to excavate become particularly salient when we look at fundamentalism. It appears that *religious* fundamentalism undermines explanatory coexistence while *magical* fundamentalism does not. There is little doubt that a Christian fundamentalist will embrace creationism and reject evolution wholeheartedly. On the other hand, the Zande “witchcraft fundamentalist” who believes in spells and witches beyond doubt will still acknowledge without demur that when a granary collapses it is not only because of the witch’s agency but also because of termites and physical laws. The non-competitiveness of magic and science seems to be bidirectional: it is hard to conceive of a system of magic in which supernatural explanations have entirely ruled out natural ones; and conversely, it is hard to conceive of a secularization process which would have undermined every magical belief.

Legare et al. (2012) have proposed that three types of explanatory coexistence should be identified.¹⁶ Following the previous remarks, we suggest that two additional types of coexistence should be recognized: *forced coexistence* and *easy coexistence*. Cases where magical explanations coexist with scientific ones illustrate what we have in mind when we speak of easy coexistence. In situations such as those described by Case 1 (Bocage witchcraft) and Case 2 (Zande witchcraft), beliefs in magic on the one hand and in biology and physics on the other are combined rather smoothly. Religious explanations and scientific ones, on the other hand, are more difficult to put together. As illustrated by the attempts of progressive popes to make biblical accounts of the world compatible with those of the most

(Footnote 15 continued)

demystification attempts against magic do not seem to be as successful as those against religion. In this regard, it is worth noting that a series of recent studies seem to invalidate Gilovich et al.’s analysis and to demonstrate that the hot hand phenomenon is in fact real (Bocskocsky, Ezekowitz, & Stein, 2014; Raab, Gula, & Gigerenzer, 2012; Sun & Wang, 2010). By and large, when a cognitive psychologist objects to believers in magic that their beliefs are demystified by such and such bias in probabilistic reasoning, believers can still reply: “the existence of a cognitive bias in general does not prove that the very specific complexity drop which happened to me was the result of that bias rather than the manifestation of a genuine non-random (i.e., intentional) process.” Magic aims to explain singular events (not events in general), and science can hardly challenge it in this regard.

¹⁶The first type of coexistence involves an explanatory pluralism across situations or contexts (it is a case of explanatory coexistence only *lato sensu* because coexistence is not considered under the same conditions). Namely, depending on context, people are likely to explain the same phenomenon (e.g., death) either in natural or in supernatural terms. The second type of coexistence consists in explaining the same phenomenon by resorting to two distinct explanatory frameworks and letting it rather vague as to how the two frameworks are exactly working together. In the third type of coexistence, two frameworks are used to explain distinct aspects of the phenomenon at hand; the explanatory function of each framework is thus clearly specified.

advanced science, a lot of amendments and hybridizations are required to ensure explanatory coexistence between science and religion.

Shtulman and Lombrozo (2016) have thoroughly examined the different ways in which explanatory coexistence can be accounted for. They find it rather unconvincing to argue that supernatural explanations persist—and coexist with natural ones—because people still lack proper scientific knowledge or because intuitive ontologies (core knowledge) make them think unscientifically, or because system 1 (fast unreflective thinking) is not able to comply with the reflective products of system 2 (slow reflective thinking). Shtulman and Lombrozo reject all these accounts and advocate a fourth explanation contending that supernatural thinking is still there simply because of its utility (2016, pp. 59–60). They further suggest that understanding the roots of explanatory coexistence probably requires a multifarious account. Here again, it seems that clearly distinguishing between magic and religion could help us better solve the challenge at hand. We have argued that humans have a craving for accounts that can explain complexity drops. Magic is extremely good at fulfilling this craving. So, as far as the coexistence of magical and natural explanations is concerned, Shtulman and Lombrozo's thesis that coexistence is underlain by utility seems quite convincing. However, it might be argued that some other account of explanatory coexistence is to be found when it comes to the coexistence of religion and science.

Notwithstanding all this, it is important to acknowledge that not any magical account of complexity drop is equally immune to scientific findings. Specifically, it seems that in the three case studies we have examined, only two clearly remain independent from science: namely, when magical explanations are triggered by an abnormally simple sequence (Type 1 complexity drop) or by abnormally simple characters and locations (Type 2 complexity drop). By contrast, it appears that when magical explanations are triggered by an atypical morphology or behavior (Type 3 complexity drop), science remains liable to undermine such explanations. In this case, magical explanations and natural explanations seem to be in competition and the coexistence between the two will be forced rather than easy. Let us consider the example of an albino squirrel which is purported to have supernatural powers because of its atypical features. If Siberian shamans were taught about genetics and about the biological underpinnings of albinism, it is likely that they would discard their belief about the supernaturalness of albino squirrels; at least, hard work would be needed in order to make natural explanations (the genetic account of albinism) and supernatural explanations (the shamanistic account of albinism) coexist. While scientific explanations cannot easily demystify magic based on Type 1 and Type 2 complexity drops, it appears that they can readily do so with magic based on Type 3 complexity drops. When assessing the impact of science and secularization on magic, one should therefore always pay close attention to the type of complexity drop at work.

Developmental Psychology and the Anthropological Challenge

Throughout this chapter, we have introduced a new model of magic in which probabilistic reasoning plays a key role. This model contrasts with existing psychological accounts of children's concepts of magic in which the violation of domain-specific intuitions has been the main focus. However, it must be emphasized that the apparent conflict between these two models of magic should not be understood as the conflict between two distinct scientific camps trying to model a single cognitive mechanism but rather as two complementary approaches researching two distinct cognitive mechanisms.

The first part of this chapter reviewed empirical studies investigating the "counterintuitive" model of magic. Children seem to perceive and understand the occurrence of events violating their domain-specific knowledge as magic. By contrast, anthropological reports about magic better fit with the "probabilistic" (or "algorithmic") model of magic. It appears that humans are liable to deploy distinct concepts of magic and therefore that magical beliefs and practices stem from two different sets of cognitive mechanisms.

It could be tempting to have a developmental story as to how "counterintuitive-magic" gradually transmutes into "probabilistic-magic", as people grow older. Such a story would allow developmental psychologists to take the anthropological challenge not too seriously: the relevance of the probabilistic model of magic would be restricted to adulthood. Until proven otherwise, however, there is no evidence to borne out such a developmental story. Hence the first challenge: *as yet, developmental psychologists have been exclusively studying magic through the lenses of the counterintuitive model of magic and they have disregarded the possibility that magical beliefs and practices are extensively underpinned by probabilistic mechanisms, not only among adults, but also among children.*

Admittedly, a few psychologists have already paid some attention to the relationship between impossibility, improbability, and magic. Such is the case of Shtulman and Carey (2007). One significant finding reported in this study is that children have it that an impossible event can occur in the real world only thanks to some magical interference, whereas an improbable event can naturally occur in the real world (2007, pp. 1026–1027). At first sight, this seems to provide evidence against the first challenge just outlined: as far as children are concerned, it is wrong to claim that improbable events trigger magical explanations. It is worth looking at the definition of "improbability" Shtulman and Carey operationalize in their experiment, though. The kind of event they have in mind when speaking of "improbable events" are events such as "finding an alligator under the bed," importantly, in their experiment, they explicitly rule out any "statistical" definition of improbable events (2007, pp. 1017–1018). As we have seen throughout the chapter, the probabilistic model of magic predicts that only events involving some kind of complexity drop are likely to trigger magical explanations. While statistical improbability does generate such complexity drops, on the other hand, "conceptual

improbability” (e.g., finding an alligator under the bed) notably fails to do so. As a consequence, Shtulman and Carey’s study does not bear on the definition of probabilistic-magic that we are proposing here. To our knowledge, psychologists have never studied the link between complexity drops and magical explanation. The challenge anthropology is posing to psychology thus seems to remain unaddressed.

In order to properly address this challenge, psychologists should introduce new experimental paradigms in their toolbox. Experimental paradigms usually used by psychologists involve children who are typically presented with events violating intuitive ontologies (e.g., a flying physical object) and a character who is shown to be able to bring about such counterintuitive events. By contrast, the kind of paradigm satisfactorily operationalizing the probabilistic definition of magic advocated here would feature the occurrence of striking complexity drops (e.g., drawing a white ball from an urn containing mostly black balls) and a character shown to be able to bring about such counterintuitive events. Indeed, from our probabilistic standpoint, a supernatural agent (or a human endowed with supernatural powers) is nothing more than an agent able to bring about complexity drops left unexplained by ordinary models of causation.

Not only does the complexity drop model suggest that psychologists may have forgotten to take into consideration the probabilistic mechanisms underlying magical explanations, but, what is more, it may be that these probabilistic mechanisms are in fact much more used and widespread than the domain-specific ones on which psychologists are focusing. Classical experimental paradigms of magic consist in presenting children with events violating intuitive ontologies. It is undeniable that children tend to resort to magical explanations when faced with such counterintuitive stimuli. Such stimuli are massively produced by psychologists in laboratories or conjurers on stages in theaters, but it is worth noting that these stimuli are very rare in the natural world. If one endorses a naturalistic approach and rejects the existence of miracles, one must conclude that counterintuitive stimuli exist only to the extent that humans use their tricks and ingenuity to produce them. On the other hand, stimuli characterized by a complexity drop do not always require the interference of human agency. A succession of unfortunate events sometimes occurs naturally whereas a miracle always requires the interference of some kind of trickery.

Contrary to what authors such as Le Bouyer de Fontenelle (1687) mistakenly asserted, a modicum of anthropological knowledge suffices to realize that agents supposedly endowed with supernatural powers (witches, sorcerers, shamans, etc.) are not claiming to be able to bring about impossible events but rather to bring about complexity drops. As a consequence, being exposed to a culture where magical practices are pervasive does not amount to being exposed to numerous counterintuitive events. It seems fair to say, then, that both children and adults are presented in their daily existence with a substantial amount of naturally occurring complexity drops while counterintuitive events remain rather rare. Therefore, the worry is that, when presenting children with counterintuitive events, psychologists do target a cognitive faculty children have (that of interpreting such events magically) but they may fail to specifically target the cognitive mechanisms children are

tapping into when they are faced with complexity drops in their daily existence. Hence, the second challenge that anthropology poses to psychology: *as yet, psychologists have been exclusively studying magical explanations triggered by counterintuitive stimuli and it may be that such magical explanations remain in fact rarely employed in the real world because of the lack of naturally occurring counterintuitive stimuli; conversely, it may be that magical explanations triggered by complexity drops are the only ones massively employed in the real world—indeed, the natural environment is replete with complexity drops while counterintuitive events are to be encountered mainly in psychology laboratories and on stages.*

Counterintuitive-Magic and Probabilistic-Magic Across Development

The comparative examination of psychological and anthropological studies on magic led us to single out two types of magic: (1) *counterintuitive-magic* (for short, *CT-magic*) is characteristically triggered by counterintuitive stimuli and is underpinned by domain-specific cognitive mechanisms; (2) *probabilistic-magic* (for short, *PR-magic*) is characteristically triggered by complexity drops and is underpinned by probabilistic (or algorithmic) cognitive mechanisms. Different hypotheses can be put forward as to how these two types of magic coexist together, or replace each other, or prevail over the other across development.

Four Hypotheses About the Developmental Trajectory of Counterintuitive-Magic and Probabilistic-Magic

In what follows, we will consider four potential hypotheses about the developmental trajectory of CT-magic and PR-magic (see summary in Table 1):

Let us spell out the content of Table 1:

Hypothesis 1. Replacement of CT-magic by PR-magic throughout development. The first hypothesis has it that CT-magic is the only kind of magic to be found

Table 1 “Counterintuitive-magic” and “probabilistic-magic” across development

	Childhood	Adulthood
Hypothesis 1	CT-magic	PR-magic
Hypothesis 2	CT-magic + PR-magic	PR-magic
Hypothesis 3	PR-magic	PR-magic
Hypothesis 4	CT-magic + PR-magic	PR-magic + CT-magic

among children while PR-magic is the only kind of magic to be found among adults.

Hypothesis 2. Decrease of CT-magic after short coexistence of PR- and CT-magic. The second hypothesis advances that children are capable both CT- and PR-magic, but that, as they grow older, they eventually discard CT-magic and keep exclusively PR-magic.

Hypothesis 3. Exclusivity of PR-magic throughout development. The third view claims that both children and adults exert exclusively PR-magic.

Hypothesis 4. The coexistence of CT-magic and PR-magic throughout development. The fourth hypothesis states that CT-magic and PR-magic coexist both among children and adults.

Hypothesis 1: Replacement of Counterintuitive-Magic by Probabilistic-Magic Throughout Development

The strongest evidence in favor of Hypothesis 1 is that, on the one hand, plenty of findings from developmental psychology support the existence of CT-magic, and that, on the other hand, plenty of data from anthropology support the existence of PR-magic. Clearly, the most straightforward way to solve these two sets of findings is to advance that developmental psychologists accurately account for magic at work *in children's minds* while anthropologists accurately account for magic at work *in adults' minds*.

There are some reasons to reject Hypothesis 1, though. As it happens, we have recently designed an experiment operationalizing the probabilistic definition of magic developed in this chapter,¹⁷ and one of us (MF) has conducted this experiment in the Peruvian Amazon, with children aged from 5 to 9. Interestingly, the preliminary data suggest that at least some children are able to discriminate between

¹⁷In this experiment, the supernatural agent (the shaman, the witch, the medicine man) was defined as an agent able to control complexity drops occurring in the world. Children were presented with videos featuring a character drawing balls from three urns located in front of him. These urns contained black balls and only one single white ball. Importantly, white balls were the only balls containing a reward inside: thus, drawing a white ball was tantamount to being lucky and drawing a black ball tantamount to being unlucky. If the character was drawing a black ball, he was being *expectedly* “unlucky” (expectedly, because most of the balls were black in the box); if, on the other hand, he was drawing a white ball, he was being *unexpectedly* “lucky” (unexpectedly, because only a single ball was white in the box). Because urns contained mostly black balls and because balls were being drawn randomly, the character was expected to be most of the time expectedly “unlucky”. This is precisely what was happening in the first phase of the experiment (no improbable outcome occurred). But, in the second phase, the character started to draw balls with a very special glove or performing a special ritual before drawing the balls. Thanks to this “magical” glove or to this “magical” ritual, the character was now drawing only white balls from the urns. The character was thus being unexpectedly lucky; he was somehow able to control the probabilistic unfolding of the events of the world and to make happen only fortunate events (i.e., to make happen only drawings whose outcomes were white balls).

random complexity drops and complexity drops seemingly resulting from the interference of an agent endowed with magical powers. Furthermore, when asked about the occurrence of a complexity drop, a few children explicitly reported that the complexity drop they had observed in the video had been caused by magic. This demonstrates that at least some children do resort to magical explanations when no ordinary model of causation can satisfactorily account for the occurrence of a complexity drop. In other words, PR-magic is already at work among children (at the very least, among some children). This is no big surprise given that very young children are known to be good probabilists (Denison, Reed, & Xu, 2013) and good at inferring agency from complexity drops (Kushnir et al., 2010; Ma & Xu, 2013; Wellman et al., 2016). Taken together, these results tend to invalidate Hypothesis 1.

All this being said, the claim that PR-magic is to be found already among children may need to be a bit qualified. One possibility is that children already possess PR-magic but that this kind of magic is not as widespread among children as it is among adults. A long time ago, Mead (1932) had noticed that supernatural thinking could not be boiled down to some kind of intellectual immaturity; by the same token, some experimental studies have recently showed that supernatural explanations increase with adulthood. This has been clearly demonstrated in Legare's study on supernatural explanations of illness in South Africa (Legare & Gelman, 2008). Such findings nicely echo thorough ethnographic work revealing the importance of training in the development of magical thinking (e.g., Luhrmann 1991). To put it in our own terms, detecting complexity drops in the environment may be a universal ability but it certainly requires intense training to properly develop. These remarks strongly suggest that although PR-magic is most likely already at work in young children's mind, it is not yet fully developed.

On a related note, one could surmise that some types of complexity drops (e.g., Type 1) are readily detected by everyone, including children, while other types of complexity drops (e.g., Type 2 or Type 3) require a great deal of enculturation. It is also worth emphasizing that magical inferences are amply guided by background knowledge. Explaining an abnormal complexity drop requires the possession of a theoretical framework specifying what kind of supernatural entity usually causes such and such complexity drop. Cultures in which complex cosmologies are available explain different types of complexity drops by resorting to different types of supernatural entities. For example, depending on whether the illness is short and sudden or gradual and long-lasting, it will be inferred that a witch rather than a sorcerer or a spirit rather than God, has caused the illness. Such inferences seem to require a cognitive sophistication that children may not have. Thus, even though PR-magic may already be at work among children, it remains unlikely that children are able to employ magic with as much sophistication as adults.

Hypothesis 1 and Hypothesis 2: Decrease of Counterintuitive-Magic Throughout Development

Both Hypothesis 1 and Hypothesis 2 advance that CT-magic decreases as children grow older. While younger children and older children are both able to detect the specialness of impossible events (Johnson & Harris, 1994), only the former judge that such events can still occur in real life through the interference of magic (e.g., Phelps & Woolley, 1994). A recent neuroimaging study shows that the recognition of impossible events and the subsequent reflective effort to make sense of the causal incongruity of such events coincide with an increased activity in the dorsolateral prefrontal cortex (DLPFC) and in the anterior cingulate cortex (ACC) (Parris, Kuhn, Mizon, Benattayallah, & Hodgson, 2009). This finding interestingly echoes other studies showing that prefrontal areas—and particularly the DLPFC—are notoriously deactivated in dreaming (Hobson, Pace-Schott, & Stickgold, 2000; Maquet et al., 1996; Schwartz & Maquet, 2002). It is well known that impossible events occurring in dreaming are almost never recognized as incongruous—presumably because of the deactivation of the DLPFC. Contrariwise, in lucid dreaming, where the activity of the DLPFC resembles that of waking states, impossible events are easily recognized as incongruous (Dresler et al., 2012; Schmitz, Kawahara-Baccus, & Johnson, 2004; Voss, Holzmann, Tuin, & Hobson, 2009). More to the point, neuroanatomical studies on the development of the prefrontal cortex among children allow us to speculate that the reason why children are able to recognize the specialness of impossible events and yet unable to demystify such events (to treat them as mere trickery) is that their prefrontal cortex (and in particular their DLPFC) remains largely underdeveloped (Diamond, 2002; Giedd et al., 1999). However, this hypothesis should be cautiously considered, because, first, the DLPFC seems also to be recruited when improbable (but possible) stimuli are processed, and second, the distinction between improbable and impossible stimuli may simply be a question of degree of activation (Fletcher et al., 2001; Fugelsang & Dunbar, 2005; Parris et al., 2009).

It could consequently be argued that psychological and neuroscientific findings support to a certain extent the claim (common to Hypotheses 1 and 2) that CT-magic significantly decreases throughout development. As humans grow older (as, presumably, their DLPFC and ACC develop), they become more and more skeptical about impossible events and they come to demystify impossible stimuli they are presented with. Unlike impossible events, complexity drops appear to remain unaffected by the development of prefrontal critical faculties. Indeed, complexity drops do not entail the occurrence of events violating basic causal laws. The existence of CT-magic depends on the absence of prefrontal scrutinizing of the world; on the other hand, PR-magic depends only on the ability to use complexity drops as a cue to agency (Griffiths & Tenenbaum, 2001; Kushnir et al., 2010; Ma & Xu, 2013; Wellman et al., 2016). Hypotheses 1, 2, and 3 all reason that if the development of prefrontal critical faculties coincides with the decrease of CT-magic, then, the only type of magic to be found in adulthood will be PR-magic.

Hypothesis 3: Exclusivity of Probabilistic-Magic Throughout Development

Hypothesis 3 goes even further than simply saying that only PR-magic is to be found among adults. It additionally contends that PR-magic is the only type of magic children are capable of. One motivation for endorsing this hypothesis is that, as argued earlier, it might be that CT-magic is after all nothing more than an experimental artefact to be encountered only in psychology laboratories where children are presented with stimuli violating their intuitive ontologies; in the real world, where no such naturally occurring stimuli are encountered, children would never have the chance to exert CT-magic. Therefore, in natural settings, the only kind of magic to be encountered among humans (regardless of their age) would be PR-magic.

The soundness of this Hypothesis depends on what is meant by saying that only PR-magic is *present* among children. This can be interpreted in two ways: (1) in the everyday world, children use only PR-magic; (2) children resort to magic only when they are presented with improbable stimuli and not when they are presented with impossible ones (regardless of whether this is done in a laboratory or in the everyday world). While (1) is concerned with how children think in natural settings, (2) is interested in studying children's cognitive dispositions and doing so may include the use of stimuli rarely (or even never) occurring in natural settings. By and large, it seems, anthropologists tend to find question (1) more significant and interesting than question (2), and conversely, psychologists tend to find (2) more significant than (1). We do not want to enter into this debate here, but let us simply point out that if Hypothesis 3 is understood to mean (2) it is then clearly untenable for, as we have seen, plenty of evidence demonstrates that when presented with the right stimuli children do use CT-magic. By contrast, if what is meant is no more than (1), then further investigations should be carried out in order to assess whether this hypothesis is accurate or not.

Hypothesis 4: The Coexistence of Counterintuitive-Magic and Probabilistic-Magic Throughout Development

Hypothesis 4 denies the main claim of Hypothesis 3 regardless of whether this claim is understood as meaning (1) or (2). To be sure, no naturally occurring impossible stimuli can be observed in the real world but it remains that advanced technology or illusions do generate such stimuli; consequently, it seems disputable to claim that no counterintuitive event is ever encountered outside of psychology laboratories and theater stages. As we have seen, children do not always possess the critical tools to demystify counterintuitive events (to recognize them as being no more than the product of trickery or technoscience). For this reason, like Hypotheses 1 and 2, Hypothesis 4 predicts that children often use CT-magic both

inside and outside of laboratories: counterintuitive stimuli brought about by tricks and technological achievements are seen as real magic. Hypothesis 4 is bolder than Hypotheses 1 and 2 as it claims that CT-magic is widespread no less among adults than among children. To be sure, adults have the ability to demystify impossible events; yet, Hypothesis 4 insists that even though adults are able to demystify impossible events, it does not mean that they always do so. Experimental evidence supports the view that, at an implicit level, adults to behave as if they were believing in CT-magic (Subbotsky, 2014). Moreover, it seems that when one's image of the world allows for some miracles to occur, not all impossible events are being demystified (Luhmann, 1991, 2012).

Hypothesis 4 has probably some grain of truth. What is more doubtful, however, is whether CT-magic is as important as PR-magic in adult cognition. Cultural differences seem to matter, here. While it is very hard to find examples of impossible events interpreted in magical terms in witchcraft or shamanism, it is undeniable that such examples can readily be pointed out in the Christian tradition. Let us now turn to the question of knowing how the types of magic are instantiated across cultures.

Types of Magic Across Cultures

Types and Subtypes of Magic

Different types (and subtypes) of magic have been identified throughout the chapter. The two types of magic are: PR-magic and CT-magic. Furthermore, it has been shown that CT-magic can be further divided into two subtypes of magic depending on whether the violation at work is minimal or maximal. Likewise, it has been shown that PR-magic can be further divided into three subtypes, which correspond to the three mechanisms by which complexity drops can occur. All these types and subtypes of magic are summarized in Table 2.

A few remarks are in order. As previously explained, entities exhibiting atypical features do violate knowledge, but only domain-general knowledge, not domain-specific one. Atypicalities trigger counterintuitions but not in the modular sense of the term. Here, when we speak of CT-magic, we specifically refer to counterintuitions violating domain-specific knowledge. This is why violations of domain-general knowledge coinciding with complexity drops are classified as PR-magic rather than CT-magic. When discussing the different types of counterintuitions, we identified three types of counterintuitions as well as a fourth type that we called a contradiction (e.g., “God is trine and one”). This fourth type of counterintuition is not featured in Table 2. The reason is that while counterintuitive beings and events are talked about both in magic and religion, contradictory beings and events, on the other hand, seem to belong only to religion.

The important question we now want to tackle is that of knowing whether the types and subtypes of magic summarized in Table 2 are liable to be more developed

Table 2 The types and subtypes of magic

Types of magic	Probabilistic-magic (PR-magic)			Counterintuitive-magic (CT-magic)	
Cognitive mechanisms involved	Tapping into probabilistic/algorithmic mechanisms			Tapping into modular mechanisms	
Subtypes of magic	<ul style="list-style-type: none"> • Type 1 complexity drop (simplicity stemming from redundancy) 	<ul style="list-style-type: none"> • Type 2 complexity drop (simplicity stemming from familiarity/proximity) 	<ul style="list-style-type: none"> • Type 3 complexity drop (simplicity stemming from atypicality) • Minimal violation of domain-general knowledge 	<ul style="list-style-type: none"> • Minimal violation of domain-specific knowledge 	<ul style="list-style-type: none"> • Maximal violation of domain-specific knowledge
Examples	Favret-Saada (1980) Brown (1986)	Evans-Pritchard (1976)	Stépanoff (2015) Keifenheim (1999)	Boyer (2001)	Atran (2002)

in certain cultures than others. Specifically, two proposals are worth examining: (1) CT-magic may be more developed in Christian cultures and less so in shamanistic or witchcraft ones; (2) PR-magic based on Type 3 complexity drop seems to be rarely encountered in modern and Western cultures, whereas it seems to be widespread in indigenous and more traditional cultures.

Counterintuitive-Magic in Christian Tradition

Owing to the importance it attaches to miracles, Christian tradition differs from numerous other supernatural systems. One first answer, then, would be to say that CT-magic is virtually absent in witchcraft or shamanism but is relatively widespread in Christian cultures. As we know, the Epistles describe many miracles that Jesus is purported to have accomplished (Jesus walking on water, feeding the multitude or changing water into wine, etc.) (Pyysiäinen, 2008). Such narratives are often discussed and commented on among Christians. It is hence plausible that the view that magic is first and foremost a matter of counterintuition originates itself in the Christian concept of magic as the power to make miracles. Scholars who are not familiar with magic in non-Christian contexts should therefore be careful not to project the Christian framework onto other cultures where such a framework is irrelevant.

Probabilistic-Magic in Christian Tradition

It must be emphasized that the importance of CT-magic in Christian culture is not necessarily incompatible with the pervasiveness of PR-magic. What anthropologists of Christianity have precisely shown is that, in their daily life, Christian believers use mostly PR-magic (Wilson, 1983). It is worthy to point out that most of Christian believers have never experienced miracles. Complexity drops, on the contrary, are part of everybody's existence. Qua humans, Christians believers are like everybody else: they mainly have to monitor and control complexity drops (improbable events), and in this respect, miracles (impossible events) do not matter greatly. There is no much difference between an animistic Shipibo singing to the spirit of *toé* (*Brugmansia*) to recover a lost item and a Roman Catholic believer praying Saint Anthony to recover a lost item. Such daily practices are all about managing complexity drops and have not much to do with impossible events. Finally, it must be noted that the term "miracle" is sometimes used in Christian contexts to refer to improbable rather impossible events (Mosse, 2006).

Counterintuitive-Magic and Probabilistic-Magic in Shipibo Culture

Participant observation is particularly useful to assess the distribution and pervasiveness of distinct (sub)types of magic within a culture. During our stay in Shipibo

communities of the Middle Ucayali, we noticed daily talks about PR-magic. This type of magic is typically used to explain why a man is unusually lucky (or unlucky) in hunting (or fishing), why an animal behaves atypically, why a child is ill, etc. In all these cases a complexity drop is taking place and Shipibo people make sense of such a drop by resorting to the world of spirits, shamanism, and sorcery. By contrast, very rarely did we notice any talk about CT-magic. Only one clear such case has been reported to us. Leonardo (now an elder) told us that one day, while he was still a child, a mestizo magician able to transform humans into animals arrived in the village and threatened Leonardo's father to turn him into a nag if he was not to give away a pig. This story is clearly based on CT-magic; but such stories are very rare in communities of the Middle Ucayali.

On the other hand, some aspects of Shipibo thinking suggest that CT-magic may be more present in indigenous communities of the Middle Ucayali than in Western countries. Although CT-magic seems very rarely used by Shipibo people, it is not unlikely that, when presented with counterintuitive stimuli (impossible events), Shipibo adults will be more inclined than Western adults to interpret these stimuli in magical terms. Indeed, ethnographic anecdotes suggest that the reality/fantasy distinction and the ascription of unreality to impossible events (i.e., recognizing them as the mere product of trickery) greatly varies across cultures. For example, Shipibo adults do not seem to grasp the difference between present-day real beings, real beings of the past, and fictional beings.

Television has recently arrived in the indigenous communities of the Middle Ucayali, and as a result, more and more Shipibo adults have occasionally the opportunity to watch documentary films and movies. To our surprise, we were once asked the following by a Shipibo adult: "Martin, tell me, are there dinosaurs living in your country?" As it happens, our informant had recently watched *Jurassic Park*. Because he did not have the background knowledge that we have about special effects, it seems that he understood the entire movie as a nonfictional film: he reasoned that people who recorded this video had traveled somewhere on earth where these astonishing animals called dinosaurs are actually living. If you do not know about special effects, such a line of reasoning is indeed very rational. Similarly, a colleague reported us that he has had the opportunity to watch *Avatar* with his Shipibo friends; they were all fascinated and bewildered and kept asking him during the entirety of the movie what kind of animal or human Na'vi were.

On another occasion, we were trying to impress our Shipibo informant by telling him that, in Western countries, some people jump from high cliffs and fly in the sky thanks to wingsuits. To which our informant responded at once that he was already knowledgeable about it and even knew the name of the man accomplishing such feats: Superman. Indeed, not long ago, he had had the opportunity to watch *Superman* and manifestly understood Superman's feats as nonfictional facts. Taken together, this ethnographic evidence implies that the ascription of reality or unreality to fictional beings and impossible events varies across cultures.

Subtypes of Probabilistic-Magic Across Cultures

It appears that subtypes of magic are also unevenly distributed across cultures. For example, while Type 3 complexity drops are encountered in Siberian and Amerindian indigenous cultures (Keifenheim, 1999; Stépanoff, 2015), it is hard to find any instance of it in modern magic (Blumberg, 1963; Favret-Saada, 1980). This may be related to the point we made earlier to the effect that Type 1 and Type 2 complexity drops, unlike Type 3, are hard to demystify and can thereby overcome secularization. Another potential cultural specificity must be pointed out. It has been mentioned that Type 3 complexity drop can take at least two forms: morphological and behavioral. Now, an interesting hypothesis worth exploring is that in cultures in which livestock farming prevails atypical features interpreted in magical terms are chiefly the color, shape, and size of animals whereas in cultures in which hunting prevails atypical features interpreted in magical terms are chiefly the behavior, smell and sound of animals. The comparison of Southern Siberia on the one hand (Stépanoff, 2015) and Northern Siberia and Amazonia on the other (Déléage, 2005; Gutierrez Choquevilca, 2010; Keifenheim, 1999; Willerslev, 2007) is in this regard very enlightening.

Conclusion

We have first presented the widely adopted view in psychology according to which magic is underpinned by domain-specific mechanisms: when faced with impossible events, young children—and some old children and adults—are inclined to resort to magical explanations. By contrast, anthropological data suggest that, in the wild, magic (i.e., shamanism, witchcraft, sorcery, etc.) is usually underpinned by domain-general mechanisms: when faced with complexity drops, adults—and, most likely, children—are inclined to resort to magical explanations. We have argued that these two models of magic are not to be understood as two accounts of the same phenomenon but rather as two accounts of distinct cognitive mechanisms.

One could object that anthropological findings fall out of the scope of psychology: what matters is not how people think in their everyday life but how they perform if asked to carry out an experimental task. When comparing the psychological and the anthropological approach to magic, it is therefore always important to bear in mind the distinction between *cognitive dispositions* (psychology's scope) and *actual cognitive performances* (anthropology's scope). Nevertheless, we have argued that anthropological data do pose a challenge to the psychological model of magic. First, psychologists have not empirically tested how children react to the occurrence of complexity drops. In particular, the question of knowing whether magical explanations are more easily triggered by complexity drops or by impossible events remains to be investigated. Second, psychologists have shown that magical explanations (triggered by impossible events) tend to decrease as children grow older and as impossible events are gradually construed as trickery rather than

real magic. This piece of evidence suggests that if any kind of magical thinking is to be found among old children and adults, it will be PR-magic (i.e., magic triggered by complexity drops). It is indeed much harder to demystify magical interpretations of complexity drops than magical interpretations of impossible events. Likewise, secularization seems to undermine mostly the CT-magic and not so much the PR-magic. Magic based on probabilistic reasoning is largely compatible with science and is probably here to stay.

We hope that this new probabilistic model of magic will lead psychologists, anthropologists, cognitive scientists of religion, and historians of religion alike to investigate many of the questions and challenges which have been raised in this chapter but which will remain in abeyance until more research is carried out.

Acknowledgements We wish to thank Claire Zedelius for her detailed and helpful feedback on previous versions of this chapter. This research has been supported by an ERC Advanced Grant “Dividnorm” # 269616.

References

- Atran, S. (2002). *In Gods we trust: The evolutionary landscape of religion*. New York: Oxford University Press.
- Atran, S., & Medin, D. (2008). *The native mind and the cultural construction of nature*. Cambridge, MA: MIT Press.
- Barkow, J., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford: Oxford University Press.
- Barrett, C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, 113(3), 628–647. doi:[10.1037/0033-295X.113.3.628](https://doi.org/10.1037/0033-295X.113.3.628)
- Barrett, J. (2004a). Counterfactuality in counterintuitive religious concepts. *Behavioral and Brain Sciences*, 27, 731–732.
- Barrett, J. (2004b). *Why would anyone believe in God?*. Walnut Creek, CA: AltaMira Press.
- Baumard, N., Hyafil, A., Morris, I., & Boyer, P. (2015). Increased affluence explains the emergence of ascetic wisdoms and moralizing religions. *Current Biology: CB*, 25(1), 10–15. doi:[10.1016/j.cub.2014.10.063](https://doi.org/10.1016/j.cub.2014.10.063)
- Benedict, R. (1933). Magic. In E. Seligman (Ed.), *Encyclopaedia of the social sciences* (pp. 39–44). New York: Mcmillan.
- Bering, J., & Parker, B. (2006). Children’s attributions of intentions to an invisible agent. *Developmental Psychology*, 42(2), 253–262. doi:[10.1037/0012-1649.42.2.253](https://doi.org/10.1037/0012-1649.42.2.253)
- Bersabé, R., & Martínez Arias, R. (2000). Superstition in gambling. *Psychology in Spain*, 4(1), 28–34.
- Bleak, J., & Frederick, C. (1998). Superstitious behavior in sport: Levels of effectiveness and determinants of use in three collegiate sports. *Journal of Sport Behavior*, 21(1), 1–15.
- Blumberg, P. (1963). Magic in the modern world. *Sociology and Social Research*, 47, 147–160.
- Bocskocsky, A., Ezekowitz, J., & Stein, C. (2014). The hot hand: A new approach to an old “fallacy.” In *8th Annual MIT Sloan Sports Analytics Conference*. Boston, MA.
- Bonnay, D., & Laugier, S. (2003). La logique sauvage de Quine à Lévi-Strauss. *Archives de Philosophie*, 66(1), 49–72.
- Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.

- Boyer, P., & Barrett, C. (2005). Domain specificity and intuitive ontology. In D. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 96–118). Hoboken, NJ: Wiley.
- Bronner, G. (2003). *L'empire des croyances*. Paris: Presses Universitaires de France.
- Bronner, G. (2007). *Coincidence: Nos représentations du hasard*. Paris: Vuibert.
- Brown, M. (1986). *Tsewa's gift: Magic and meaning in an Amazonian society*. Tuscaloosa: University of Alabama Press.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Chandler, M., & Lalonde, C. (1994). Surprising, magical and miraculous turns of events: Children's reactions to violations of their early theories of mind and matter. *British Journal of Developmental Psychology*, 12(1), 83–95. doi:[10.1111/j.2044-835X.1994.tb00620.x](https://doi.org/10.1111/j.2044-835X.1994.tb00620.x)
- Chater, N., & Vitanyi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences*, 7(1), 19–22.
- Clément, F. (2003). L'esprit ensorcelé: Les racines cognitives de la sorcellerie. *Terrain*, (41), 121–136. doi:[10.4000/terrain.1670](https://doi.org/10.4000/terrain.1670)
- Cook, C., & Sobel, D. (2011). Children's beliefs about the fantasy/reality status of hypothesized machines. *Developmental Science*, 14(1), 1–8. doi:[10.1111/j.1467-7687.2009.00949.x](https://doi.org/10.1111/j.1467-7687.2009.00949.x)
- Da Costa, N., Bueno, O., & French, S. (1998). Is there a Zande logic? *History and Philosophy of Logic*, 19(1), 41–54. doi:[10.1080/01445349808837295](https://doi.org/10.1080/01445349808837295)
- Déléage, P. (2005). *Le Chamanisme sharanahua. Enquête sur l'apprentissage et l'épistémologie d'un rituel* (PhD dissertation). École des Hautes Études en Sciences Sociales, Paris.
- Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology*, 49(2), 243–249.
- Dessalles, J.-L. (2007). Spontaneous assessment of complexity in the selection of events. In Technical Report (Ed.), *ParisTech*. Paris: Ecole Nationale Supérieure des Télécommunications.
- Dessalles, J.-L. (2010). Emotion in good luck and bad luck: Predictions from simplicity theory. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Portland: Cognitive Science Society.
- Dessalles, J.-L. (2013). Algorithmic Simplicity and Relevance. In D. Dowe (Ed.), *Algorithmic probability and friends: Bayesian prediction and artificial intelligence* (pp. 119–130). Berlin/Heidelberg: Springer.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. Stuss & R. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). New York: Oxford University Press.
- Dresler, M., Wehrle, R., Spoormaker, V. I., Koch, S. P., Holsboer, F., Steiger, A., ... Czisch, M. (2012). Neural correlates of dream lucidity obtained from contrasting lucid versus non-lucid REM sleep: A combined EEG/fMRI case study. *Sleep*, 35(7), 1017–1020. doi:[10.5665/sleep.1974](https://doi.org/10.5665/sleep.1974)
- Dweck, C. (2006). *Mindset: The new psychology of success*. New York: Random House.
- Evans, M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, 42(3), 217–266. doi:[10.1006/cogp.2001.0749](https://doi.org/10.1006/cogp.2001.0749)
- Evans-Pritchard, E. (1976). *Witchcraft, oracles, and magic among the Azande* (Abridged ed.). Oxford: Clarendon Press.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104(2), 301–318.
- Favret-Saada, J. (1980). *Deadly words: Witchcraft in the Bocage*. Cambridge/Paris: Cambridge University Press/Éditions de la Maison des Sciences de l'Homme.
- Favret-Saada, J. (1986). L'invention d'une thérapie: La sorcellerie bocaine 1887–1970. *Le Débat*, 3(40), 29–46.
- Favret-Saada, J. (1990). Être affecté. *Gradhiva*, 8, 3–9.
- Fletcher, P., Anderson, J. M., Shanks, D., Honey, R., Carpenter, T. A., Donovan, T., ... Bullmore, E. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nature Neuroscience*, 4(10), 1043–1048. doi:[10.1038/nn733](https://doi.org/10.1038/nn733)

- Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fortier, M. (in press). Le sens de réalité dans les expériences psychotropes : Étude comparée des hallucinogènes sérotoninergiques et des hallucinogènes anticholinergiques. In S. Baud (Ed.), *Des plantes psychotropes : Initiations, thérapies et quêtes de soi* (Vol. 2). Paris: Imago.
- Fortier, M. (in preparation). Supernatural thinking as algorithmic reasoning.
- Frazer, J. G. (1922). *The golden bough* (Abridged ed.). London: Macmillan.
- Fugelsang, J., & Dunbar, K. (2005). Brain-based mechanisms underlying complex causal thinking. *Neuropsychologia*, 43(8), 1204–1213. doi:10.1016/j.neuropsychologia.2004.10.012
- Furnham, A. (1988). *Lay theories: Everyday understanding of problems in the social sciences*. Oxford: Pergamon Press.
- Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F., Liu, H., Zijdenbos, A., ... Rapoport, J. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2(10), 861–863. doi:10.1038/13158
- Gille, B. (2014). Une femme, deux visages, un serpent, deux têtes: Cognition, contre-intuition et représentations chimériques. *L'Homme*, 209, 39–67.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314. doi:10.1016/0010-0285(85)90010-6
- Gopnik, A., Wellman, H., & Kuhl, P. (1999). *The scientist in the crib: What early learning tells us about the mind*. New York: William Morrow.
- Goulet, J.-G., & Miller, B. (Eds.). (2007). *Extraordinary anthropology: Transformations in the field*. Lincoln/London: University of Nebraska Press.
- Griffiths, T., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences*, 14(8), 357–364. doi:10.1016/j.tics.2010.05.004
- Griffiths, T., & Tenenbaum, J. (2001). Randomness and coincidences: Reconciling intuition and probability theory. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*.
- Griffiths, T., & Tenenbaum, J. (2003). Probability, algorithmic complexity, and subjective randomness. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*.
- Guthrie, S. (1993). *Faces in the clouds: A new theory of religion*. New York: Oxford University Press.
- Gutierrez Choquevilca, A.-L. (2010). Imaginaire acoustique et apprentissage d'une ontologie animiste: Le cas des Quechua d'Amazonie péruvienne. *Ateliers D'anthropologie. Revue éditée Par Le Laboratoire D'ethnologie et de Sociologie Comparative*, 34. doi:10.4000/ateliers.8553
- Harris, P., Brown, E., Marriott, C., Whittall, S., & Harmer, S. (1991). Monsters, ghosts and witches: Testing the limits of the fantasy—Reality distinction in young children. *British Journal of Developmental Psychology*, 9(1), 105–123. doi:10.1111/j.2044-835X.1991.tb00865.x
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57, 243–259. doi:10.2307/1416950
- Hess, P., & Allen, P. (2008). *Catholicism and science*. Westport CT/London: Greenwood Press.
- Hobson, A., Pace-Schott, E., & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, 23(6), 793–1121.
- Hollis, M., & Lukes, S. (Eds.). (1982). *Rationality and relativism*. Cambridge MA/London: MIT Press.
- Horton, R. (1967a). African traditional thought and western science. Part I: From tradition to science. *Africa: Journal of the International African Institute*, 37(1), 50–71.
- Horton, R. (1967b). African traditional thought and western science. Part II: The “closed” and “open” predicaments. *Africa: Journal of the International African Institute*, 37(2), 155–187.
- Hubert, H., & Mauss, M. (1903). Esquisse d'une théorie générale de la magie. *L'Année Sociologique*, 7, 1–146.

- Jarvie, I. C., & Agassi, J. (1967). The problem of the rationality of magic. *The British Journal of Sociology*, 18, 55–74. doi:[10.2307/588590](https://doi.org/10.2307/588590)
- Johnson, C., & Harris, P. (1994). Magic: Special but not excluded. *British Journal of Developmental Psychology*, 12(1), 35–51. doi:[10.1111/j.2044-835X.1994.tb00617.x](https://doi.org/10.1111/j.2044-835X.1994.tb00617.x)
- Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development*, 69(3), 706–720. doi:[10.1111/j.1467-8624.1998.tb06238.x](https://doi.org/10.1111/j.1467-8624.1998.tb06238.x)
- Keifenheim, B. (1999). Concepts of perception, visual practice, and pattern art among the Cashinahua Indians (Peruvian Amazon area). *Visual Anthropology*, 12(1), 27–48. doi:[10.1080/08949468.1999.9966766](https://doi.org/10.1080/08949468.1999.9966766)
- Keil, F., & Newman, G. (2015). Order, order everywhere, and only an agent to think: The cognitive compulsion to infer intentional agents. *Mind and Language*, 30(2), 117–139.
- Keinan, G. (1994). Effects of stress and tolerance of ambiguity on magical thinking. *Journal of Personality and Social Psychology*, 67(1), 48–55. doi:[10.1037/0022-3514.67.1.48](https://doi.org/10.1037/0022-3514.67.1.48)
- Kim, S., & Harris, P. (2014a). Belief in magic predicts children's selective trust in informants. *Journal of Cognition and Development*, 15(2), 181–196.
- Kim, S., & Harris, P. (2014b). Children prefer to learn from mind-readers. *British Journal of Developmental Psychology*, 32(4), 375–387. doi:[10.1111/bjdp.12044](https://doi.org/10.1111/bjdp.12044)
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96(4), 674–689. doi:[10.1037/0033-295X.96.4.674](https://doi.org/10.1037/0033-295X.96.4.674)
- Kushnir, T., Xu, F., & Wellman, H. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, 21(8), 1134–1140. doi:[10.1177/0956797610376652](https://doi.org/10.1177/0956797610376652)
- Lane, J., Wellman, H., & Evans, M. (2010). Children's understanding of ordinary and extraordinary minds. *Child Development*, 81(5), 1475–1489. doi:[10.1111/j.1467-8624.2010.01486.x](https://doi.org/10.1111/j.1467-8624.2010.01486.x)
- Lane, J., Wellman, H., & Evans, M. (2012). Sociocultural input facilitates children's developing understanding of extraordinary minds. *Child Development*, 83(3), 1007–1021. doi:[10.1111/j.1467-8624.2012.01741.x](https://doi.org/10.1111/j.1467-8624.2012.01741.x)
- Lane, J., Wellman, H., & Evans, M. (2014). Approaching an understanding of omniscience from the preschool years to early adulthood. *Developmental Psychology*, 50(10), 2380–2392. doi:[10.1037/a0037715](https://doi.org/10.1037/a0037715)
- Le Bouyer de Fontenelle, B. (1687). *Histoire des oracles*. Paris: G. de Luyne.
- Legare, C., Evans, M., Rosengren, K., & Harris, P. (2012). The coexistence of natural and supernatural explanations across cultures and development. *Child Development*, 83(3), 779–793. doi:[10.1111/j.1467-8624.2012.01743.x](https://doi.org/10.1111/j.1467-8624.2012.01743.x)
- Legare, C., & Gelman, S. (2008). Bewitchment, biology, or both: The co-existence of natural and supernatural explanatory frameworks across development. *Cognitive Science*, 32(4), 607–642. doi:[10.1080/03640210802066766](https://doi.org/10.1080/03640210802066766)
- Lévy-Bruhl, L. (1923). *Primitive mentality*. (L. Clare, Trans.). Oxford: Clarendon Press.
- Lévy-Bruhl, L. (1927). *The "soul" of the primitive*. (L. Clare, Trans.). London: Allen and Unwin Ltd.
- Lévy-Bruhl, L. (1966). *How natives think*. New York: Washington Square Press.
- Li, M., & Vitanyi, P. (1997). *An introduction to Kolmogorov complexity and its applications* (2d ed.). New York: Springer.
- Luhrmann, T. (1991). *Persuasions of the witch's craft: Ritual magic in contemporary England*. Cambridge MA: Harvard University Press.
- Luhrmann, T. (2012). *When God talks back: Understanding the American evangelical relationship with God*. New York: Alfred Knopf.
- Malinowski, B. (1948). *Magic, science and religion and other essays*. Glencoe, IL: The Free Press.
- Ma, L., & Xu, F. (2013). Preverbal infants infer intentional agents from the perception of regularity. *Developmental Psychology*, 49(7), 1330–1337.
- Maquet, P., Péters, J.-M., Aerts, J., Delfiore, G., Degueldre, C., Luxen, A., & Franck, G. (1996). Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature*, 383(6596), 163–166. doi:[10.1038/383163a0](https://doi.org/10.1038/383163a0)

- Markus, H., & Hamedani, M. (2007). Sociocultural psychology: The dynamic interdependence among self systems and social systems. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 3–39). New York/London: Guilford Press.
- Mead, M. (1932). An investigation of the thought of primitive children, with special reference to animism. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 62.
- Mosse, D. (2006). Possession and confession: Affliction and sacred power in colonial and contemporary Catholic South India. In F. Cannell (Ed.), *The anthropology of Christianity* (pp. 99–133). Durham, NC: Duke University Press.
- Nader, L. (Ed.). (1996). *Naked science: Anthropological inquiry into boundaries, power, and knowledge*. New York/London: Routledge.
- Newman, G., Keil, F., Kuhlmeier, V., & Wynn, K. (2010). Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences*, 107(40), 17140–17145. doi:10.1073/pnas.0914056107
- Paris, B., Kuhn, G., Mizon, G., Benattayallah, A., & Hodgson, T. L. (2009). Imaging the impossible: An fMRI study of impossible causal relationships in magic tricks. *Neuroimage*, 45(3), 1033–1039. doi:10.1016/j.neuroimage.2008.12.036
- Pégard, C. (2000, July 7). Mitterrand et ses magiciens. *Le Point*. Retrieved from <http://www.lepoint.fr/actualites-politique/2007-01-16/mitterrand-et-ses-magiciens/917/0/3588>
- Phelps, K., & Woolley, J. (1994). The form and function of young children's magical beliefs. *Developmental Psychology*, 30(3), 385–394. doi:10.1037/0012-1649.30.3.385
- Piaget, J. (1929). *The Child's conception of the World*. London: Routledge & Kegan Paul.
- Pyysiäinen, I. (2002). Religion and the counter-intuitive. In I. Pyysiäinen & V. Anttonen (Eds.), *Current approaches in the cognitive science of religion* (pp. 3–33). London: Continuum.
- Pyysiäinen, I. (2008). The enduring fascination of miracles. In H. Ellens (Ed.), *Miracles: God, science, and psychology in the paranormal* (Vol. 2, pp. 17–26). Westport CT: Greenwood/Praeger.
- Raab, M., Gula, B., & Gigerenzer, G. (2012). The hot hand exists in volleyball and is used for allocation decisions. *Journal of Experimental Psychology: Applied*, 18(1), 81–94. doi:10.1037/a0025951
- Reber, R., Fazendeiro, T., & Winkielman, P. (2002). Processing fluency as the source of experiences at the fringe of consciousness. *Psyche: An Interdisciplinary Journal of Research on Consciousness*, 8(10).
- Rosengren, K., & Hickling, A. (1994). Seeing is believing: Children's explanations of commonplace, magical, and extraordinary transformations. *Child Development*, 65(6), 1605–1626. doi:10.1111/j.1467-8624.1994.tb00838.x
- Rosengren, K., Johnson, C., & Harris, P. (2000). Preface. In K. Rosengren, C. Johnson, & P. Harris (Eds.), *Imagining the impossible: Magical, scientific, and religious thinking in children* (pp. xiii–xx). Cambridge/New York: Cambridge University Press.
- Rosengren, K., Kalish, C., Hickling, A., & Gelman, S. (1994). Exploring the relation between preschool children's magical beliefs and causal thinking. *British Journal of Developmental Psychology*, 12(1), 69–82. doi:10.1111/j.2044-835X.1994.tb00619.x
- Rudski, J., & Edwards, A. (2007). Malinowski goes to college: Factors influencing students' use of ritual and superstition. *The Journal of General Psychology*, 134(4), 389–403. doi:10.3200/GENP.134.4.389-404
- Rutherford, M., & Kuhlmeier, V. (Eds.). (2013). *Social perception: Detection and interpretation of animacy, agency, and intention*. Cambridge, MA: MIT Press.
- Saillenfest, A., & Dessalles, J.-L. (2015). Some probability judgments may rely on complexity assessments. *Proceedings of the 37th annual meeting of the Cognitive Science Society* (pp. 2069–2074). Austin: Cognitive Science Society.
- Schmitz, T., Kawahara-Baccus, T., & Johnson, S. (2004). Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *NeuroImage*, 22(2), 941–947. doi:10.1016/j.neuroimage.2004.02.018
- Schwartz, S., & Maquet, P. (2002). Sleep imaging and the neuro-psychological assessment of dreams. *Trends in Cognitive Sciences*, 6(1), 23–30. doi:10.1016/S1364-6613(00)01818-0

- Schwarz, N., & Clore, G. (2007). Feelings and phenomenal experiences. In A. Kruglanski & T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 385–407). New York: Guilford.
- Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Development*, 78(3), 1015–1032. doi:[10.1111/j.1467-8624.2007.01047.x](https://doi.org/10.1111/j.1467-8624.2007.01047.x)
- Shtulman, A., & Lombrozo, T. (2016). Bundles of contradiction: A coexistence view of conceptual change. In D. Barner & A. Baron (Eds.), *Core knowledge and conceptual change* (pp. 49–67). Oxford/New York: Oxford University Press.
- Spelke, E. (2000). Core knowledge. *American Psychologist*, 55(11), 1233–1243.
- Sperber, D. (1975). *Rethinking symbolism*. (A. Morton, Trans.). Cambridge: Cambridge University Press.
- Sperber, D. (1985). *On anthropological knowledge: Three essays*. Cambridge/Paris: Cambridge University Press/Éditions de la Maison des Sciences de l'Homme.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford/Malden, MA: Blackwell.
- Stépanoff, C. (2015). Transsingularity: The cognitive foundations of shamanism in Northern Asia. *Social Anthropology*, 23(2), 169–185. doi:[10.1111/1469-8676.12108](https://doi.org/10.1111/1469-8676.12108)
- Subbotsky, E. (1993). *Foundations of the mind: Children's understanding of reality*. Cambridge MA: Harvard University Press.
- Subbotsky, E. (2001). Causal explanations of events by children and adults: Can alternative causal modes coexist in one mind? *British Journal of Developmental Psychology*, 19(1), 23–45. doi:[10.1348/026151001165949](https://doi.org/10.1348/026151001165949)
- Subbotsky, E. (2004). Magical thinking in judgments of causation: Can anomalous phenomena affect ontological causal beliefs in children and adults? *British Journal of Developmental Psychology*, 22(1), 123–152. doi:[10.1348/02615100472901140](https://doi.org/10.1348/02615100472901140)
- Subbotsky, E. (2005). The permanence of mental objects: Testing magical thinking on perceived and imaginary realities. *Developmental Psychology*, 41(2), 301–318. doi:[10.1037/0012-1649.41.2.301](https://doi.org/10.1037/0012-1649.41.2.301)
- Subbotsky, E. (2010). *Magic and the mind: Mechanisms, functions, and development of magical thinking and behavior*. New York/Oxford: Oxford University Press.
- Subbotsky, E. (2014). The belief in magic in the age of science. *SAGE Open*, 4(1), 1–17. doi:[10.1177/2158244014521433](https://doi.org/10.1177/2158244014521433)
- Sun, Y., & Wang, H. (2010). Gambler's fallacy, hot hand belief, and the time of patterns. *Judgment and Decision Making*, 5(2), 124–132.
- Tambiah, S. (1990). *Magic, science, religion, and the scope of rationality*. Cambridge: Cambridge University Press.
- Tenenbaum, J., Kemp, C., Griffiths, T., & Goodman, N. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331(6022), 1279–1285. doi:[10.1126/science.1192788](https://doi.org/10.1126/science.1192788)
- Thomas, K. (1971). *Religion and the decline of magic. Studies in popular beliefs in sixteenth and seventeenth century England*. London: Weidenfeld and Nicolson.
- Triplett, T. (1994). Is there anthropological evidence that logic is culturally relative?: Remarks on Bloor, Jennings, and Evans-Pritchard. *The British Journal for the Philosophy of Science*, 45(2), 749–760.
- Unkelback, C., & Greifeneder, R. (Eds.). (2013). *The experience of thinking: How the fluency of mental processes influences*. London/New York: Psychology Press.
- Vaden, V., & Woolley, J. (2011). Does God make it real? Children's belief in religious stories from the Judeo-Christian tradition. *Child Development*, 82(4), 1120–1135. doi:[10.1111/j.1467-8624.2011.01589.x](https://doi.org/10.1111/j.1467-8624.2011.01589.x)
- Vikan, A., & Clausen, S. (1993). Freud, Piaget, or Neither? Beliefs in controlling others by wishful thinking and magical behavior in young children. *The Journal of Genetic Psychology*, 154(3), 297–314. doi:[10.1080/00221325.1993.10532183](https://doi.org/10.1080/00221325.1993.10532183)
- Vitanyi, P. M. B., & Li, M. (2000). Minimum description length induction, Bayesianism, and Kolmogorov complexity. *IEEE Transactions on Information Theory*, 46(2), 446–464. doi:[10.1109/18.825807](https://doi.org/10.1109/18.825807)

- Voss, U., Holzmann, R., Tuin, I., & Hobson, A. (2009). Lucid dreaming: A state of consciousness with features of both waking and non-lucid dreaming. *Sleep*, 32(9), 1191–1200.
- Vyse, S. (1997). *Believing in magic: The psychology of superstition*. New York: Oxford University Press.
- Weber, M. (2003). *The protestant ethic and the spirit of capitalism*. New York: Dover.
- Wellman, H., Kushnir, T., Xu, F., & Brink, K. (2016). Infants use statistical sampling to understand the psychological world. *Infancy*, 21(5), 668–676. doi:[10.1111/infa.12131](https://doi.org/10.1111/infa.12131)
- Willerslev, R. (2007). *Soul hunters: Hunting, animism, and personhood among the Siberian Yukaghirs*. Berkeley/Los Angeles: University of California Press.
- Wilson, S. (1983). *Saints and their cults: Studies in religious sociology, folklore and history*. Cambridge: Cambridge University Press.
- Woolley, J. (2000). The development of beliefs about direct mental-physical causality in imagination, magic, and religion. In K. Rosengren, C. Johnson, & P. Harris (Eds.), *Imagining the impossible: Magical, scientific, and religious thinking in children* (pp. 99–129). Cambridge: Cambridge University Press.
- Woolley, J., Boerger, E., & Markman, A. (2004). A visit from the candy witch: Factors influencing young children's belief in a novel fantastical being. *Developmental Science*, 7(4), 456–468.
- Woolley, J., & Ghossainy, M. (2013). Revisiting the fantasy-reality distinction: Children as naïve skeptics. *Child Development*, 84(5), 1496–1510. doi:[10.1111/cdev.12081](https://doi.org/10.1111/cdev.12081)
- Woolley, J., Phelps, K., Davis, D., & Mandell, D. (1999). Where theories of mind meet magic: The development of children's beliefs about wishing. *Child Development*, 70(3), 571–587. doi:[10.1111/1467-8624.00042](https://doi.org/10.1111/1467-8624.00042)
- Woolley, J., & Wellman, H. (1993). Origin and truth: Young children's understanding of imaginary mental representations. *Child Development*, 64(1), 1–17.
- Xu, F. (2007). Rational statistical inference and cognitive development. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind* (Vol. 3, pp. 199–215)., Foundations and the future New York: Oxford University Press.

Part IV
Investigation of Lay Theories About
Mental and Physical Health

Mindsets of Body Weight

Jeni L. Burnette, Crystal L. Hoyt and Kasey Orvidas

At first glance, what to believe about the nature of weight management seems straightforward—excess weight is the result of the surplus of calories consumed minus calories expended. Thus, the solution is fairly easy—eat less, move more. Although there is some truth to the idea that weight can be changed using such a simple strategy, there is also significant debate about the optimal weight loss approach. For example, some studies show that eating less is far more important than exercise for determining weight (e.g., Wilks et al., 2011), whereas others show that calorie restriction interventions inevitably fail (Mann et al., 2007; Mann, 2015). Furthermore, the public is constantly bombarded with disparate messages. For example, in the United States, the American Medical Association (AMA) labeled obesity a disease (Pollack, 2013). Does this mean weight is genetically determined? Alternatively, Michelle Obama’s “Let’s Move” campaign stresses the importance of eating healthy, getting active and taking action (Let’s Move Initiative, 2010). Does this mean people who fail to reach their weight loss goals are to blame for not changing their lifestyle?

These weight management messages highlight the fundamental issue of whether people should believe that body weight is a fixed entity or a malleable attribute. Such a lay belief or theory is also called an implicit theory or mindset (Dweck, 2000; Dweck & Leggett, 1988; Molden & Dweck, 2006). In the current chapter, we highlight the empirical research regarding the consequences of these weight-based mindsets for a range of significant outcomes—from self-regulation to fat stigma to

J.L. Burnette (✉) · K. Orvidas

Department of Psychology, North Carolina State University, 640 Poe Hall, Box 7650,
Raleigh, NC 27695, USA
e-mail: jlburne5@ncsu.edu

C.L. Hoyt

Jepson School of Leadership Studies, University of Richmond, Virginia, USA

body shame. First, however, we briefly review generally what lay theories, or mindsets, are, when they matter, and why they matter, before turning our attention to mindsets of body weight.

Mindsets Overview

What Are Mindsets?

People's mindsets are theories or beliefs that organize their world and give meaning to events in their lives. Beginning with the seminal writings of William James, consensus in the field of psychology supports the notion that personal beliefs are critical for shaping one's reality. The fundamental assumption that individuals vary in their knowledge structures, or mindsets, which in turn greatly influences attitudes and behavior is common to several perspectives on human behavior (Ross, 1989). For example, Piaget, a pioneer in developmental psychology, suggested that meaning systems are just as important as logical thinking in shaping behavior (Piaget & Garcia, 1991). Similarly, Kelly (1955) proposed, "man looks at his world through transparent templates which he creates and then attempts to fit over the realities of which the world is composed" (pp. 8–9). Continuing with these venerable traditions, Dweck's implicit theory approach suggests that beliefs often converge around two main themes: entity and incremental theories, also termed fixed and growth mindsets, respectively. The entity framework holds that a human attribute within a specific domain is unalterable. For example, in the intelligence domain, an entity theorist would hold that intelligence is a fixed trait and not much can be done to change it. In contrast, an incremental theorist believes in the malleability of human attributes and would agree more strongly that intelligence can be changed substantially through hard work and effective strategies.

In understanding generally what mindsets are, it is important to note across a variety of studies and diverse populations that (a) mindsets are generally uncorrelated with demographics and personality and (b) people can hold different mindsets in different domains (e.g., intelligence in general versus math in particular). Although it appears as if implicit theories could potentially be confounded with dispositional constructs such as optimism or self-esteem, past literature provides empirical evidence that implicit theories are not correlated with these traits, or others such as social desirability, the Big Five personality traits (i.e., extraversion, agreeableness, openness, conscientiousness, and neuroticism), or self-monitoring (Costa & McCrae, 1985, Dweck, Chiu, & Hong, 1995b). Additionally, mindsets are rarely correlated with other individual differences such as academic aptitude, education, or political stance (Dweck, Chiu, & Hong, 1995a; Niiya, Crocker, & Bartmess, 2004; Plaks & Stecher, 2007; Spinath, Spinath, Riemann, & Angleitner, 2003; Taberero & Wood, 1999). As for domain specificity, even if an individual is an incremental theorist in regard to intelligence, believing intelligence is a

malleable trait, this does not necessarily mean that those incremental beliefs will carry over into other domains (Dweck, Chiu, & Hong, 1995a). For instance, an individual can have an entity theory about moral character, but an incremental theory of intelligence. Or, alternatively, an individual might be consistently praised for his or her “natural” athletic ability and talent while practicing soccer. This particular individual then may develop a fixed theory about athletic ability. However, teachers praise this individual’s effort during academic endeavors, thereby helping the child develop an incremental theory of intelligence. In summary, implicit theories are important knowledge structures about the fixed versus malleable nature of human attributes. These beliefs or mindsets are distinct from other dispositional tendencies and are domain specific.

When Do Mindsets Matter?

Now that we understand what implicit theories or mindset are (and are not), let us turn our attention to the second question of when they matter. Dweck originally developed the implicit theory approach to explain how the two meaning systems impact persistence following setbacks (e.g., Dweck, 2000). More specifically, a primary question of interest in the literature is: Why do some individuals self-regulate and stay motivated even in the face of obstacles, whereas others throw in the towel? For example, when runners first attempted to break the 4-minute mile, what made some men continue to persist even though they were consistently advised that breaking the barrier was physiologically impossible? In the face of goal obstruction, Dweck and her colleagues proposed that a critical predictor of persistence and coping with failures is people’s implicit theories or intuitive conceptions about the stability of human traits and attributes (Molden & Dweck, 2006). Although this theoretical framework has primarily been applied to understanding motivation in the wake of academic setbacks, implicit theories have also been extended to understanding social perception, relationships, leadership, exercise, weight management, and more, with an emphasis on a desire to understand when failing inspires versus undermines motivation.

Across contexts, implicit theories predict how individuals handle ego threats—threats to the sense of self. For instance, within a math domain, an incremental theory can help buffer females against the deleterious effects of stereotype threat—feeling threatened by the potential of conforming to negative beliefs about one’s group (Burkley, Parker, Stermer, & Burkley, 2010; Steele, 1997). For example, females, even if they are anxious about confirming the negative stereotype that they are inferior to men in math, remain engaged and interested in the task if they hold a growth mindset. Furthermore, when facing ego threats, such as a setback, growth mindsets predict more effortful self-regulatory processes. For instance, in one study, regardless of whether students held an entity or incremental theory, they had similar math grades in elementary school, but when faced with the challenging transition into middle school the students’ grades began to deviate from one another based on

their implicit theory, with incremental theorists (students with growth mindsets) outperforming their fixed mindset counterparts (Blackwell, Trzesniewski, & Dweck, 2007; Dweck, 2012). Whereas incremental theorists use setbacks as information about avenues for potential improvement, entity theorists respond to such ego threats by focusing on their need to prove their abilities, viewing failures as signs pointing to their innate lack of ability. For example, students holding an entity theory of intelligence attributed a hypothetical academic failure to their inherent intellect. In contrast, students with an incremental theory attributed failure to inadequate effort (Henderson & Dweck, 1990). In summary, mindsets matter most under ego threats with growth, relative to fixed, theorists reporting greater resiliency when obstacles to goal pursuits may arise.

Why Do Mindsets Matter?

This brings us to our third question—why do mindsets matter? The fixed and growth mindset meaning systems lead people to think, feel, and act differently under similar situations. More specifically, implicit theories are an integral part of people’s motivational systems influencing self-regulatory processes and ultimately goal achievement (Burnette, O’Boyle, VanEpps, Pollack, & Finkel, 2013). Mindsets matter because they set up how individuals set, strive for, and monitor goal pursuits. Whereas individuals with growth mindsets set goals focused on learning, employ mastery-oriented strategies in the face of challenges, and report greater confidence and self-efficacy when evaluating the potential for future goal success, individuals with fixed mindsets set goals focused on performance, employ helpless-oriented strategies in the face of challenges, and report feeling vulnerable and anxious when evaluating past and future performance.

First, mindsets matter for the goals people set (Moskowitz & Grant, 2009). Growth theorists focus more on learning and fixed theorists focus more on performance. Learning goals involve striving to master a skill (e.g., Maurer, Mitchell, & Barbeite, 2002), whereas performance goals involve striving to demonstrate one’s ability, frequently relative to others (e.g., Leondari & Gialamas, 2002). For example, an athlete focused on gaining valuable skills such as mastering a header in soccer is striving for a learning-oriented goal. In contrast, an athlete focused on scoring more goals than the other team is striving for a performance-oriented goal. Second, mindsets lead people to strive for these goals in different ways. There are two central constructs relevant to mindsets and the goal striving process: mastery-oriented versus helpless-oriented strategies. Mastery-oriented strategies describe an overall “hardy response” (Dweck, 2000, p. 6), showing persistence and tenacity (Dweck, 2000) such as help seeking (e.g., Ommundsen, 2003) and practice time (e.g., Cury, Da Fonseca, Zahn, & Elliot, 2008). In contrast, helpless-oriented responses entail perceptions of lack of control and thus result in averting attention and resources away from one’s goal such as procrastinating and self-handicapping (Diener & Dweck, 1978, 1980). Individuals with growth mindsets tend to engage in

more mastery-oriented, and fewer helpless-oriented strategies when trying to reach their goals (e.g., Chen, et al., 2008; Dweck & Molden, 2005; Elliott & Dweck, 1988; Hong, Chiu, Dweck, & Wan, 1999; Nichols, White, & Price, 2006; Wang & Biddle, 2001; Wang, Chatzisarantis, Spray, & Biddle, 2002). Third, mindsets also matter as people monitor their goals. That is, when they evaluate the degree to which they are getting closer to the desired goal state and if they have the potential to achieve it in the future. Whereas individuals with fixed mindsets tend to feel stronger negative emotions, especially anxiety, individuals with growth mindsets remain confident about the potential for future success (Burnette et al., 2013).

In summary, mindsets are beliefs about whether human attributes are fixed or malleable. They matter most when individuals face ego threats or a threat to one's sense of self such as setbacks, critiques, or challenges. And, in understanding why mindsets matter, one simple answer is because they establish the framework for self-regulatory processes relevant to goal achievement.

Weight-Based Mindsets

Given the long-standing empirical support for the importance of mindsets in understanding self-regulatory processes and achievement especially when individuals face setbacks, scholars have turned to mindsets to offer insight into what leads to effective weight management strategies (e.g., Burnette, 2010; Burnette & Finkel, 2012). For example, why do two people, equally devoted to the goal of losing excess weight, so often differ in their degree of success at this task—with one managing to reach his or her dietary goals and the other giving up entirely? What helps some people continue to persist on their weight loss goals in the face of so many environmental constraints? The answer to these questions is important considering the rates of obesity (Sturm & Hattori, 2013) and the links between reaching a healthy weight and physical and psychological health (Blackburn, 1995; Pasanisi, Contaldo, De Simone, & Mancini, 2001; Rippe et al., 1998). One key to variations in persistence on diets can be found in each person's belief about weight management—their implicit theory, or mindset, about the nature of weight.

Although a number of important perspectives can help predict sustained motivation in health domains (e.g., health belief model, Maiman, Becker, Kirscht, Haefner, & Drachman, 1977; social cognitive theory, Resnicow et al., 1997; and theory of planned behavior, Ajzen, 1991; Ajzen & Fishbein, 1977; Norman, Conner, & Bell, 1999), in the current chapter, we focus on the implicit or lay theory approach. Such an approach can add to the existing literature by starting earlier in the psychological chain, by providing a mechanism to leverage for behavioral change, and by incorporating a growing body of literature supporting the importance of these mindsets for self-regulation (Burnette et al., 2013). For example, implicit theories predict many of the outcomes relevant to other cognitive perspectives such as efficacy (Bråten & Strømsø, 2005), locus of control (Dweck et al.,

1995a, b), or behavioral intentions (Dweck, Chiu, & Hong, 1995a). Furthermore, a plethora of research finds that mindsets can be shifted with scalable low-cost intervention efforts and that such efforts have high payoffs in terms of motivation and achievement (e.g., Blackwell et al., 2007; Paunesku et al., 2015). Dieting is an achievement domain with numerous parallels to previously studied areas in the mindset literature. As individuals try to manage their weight, they establish goals, experience setbacks, regulate their feelings of disappointment, and try to persist in the face of adversity. Building on seminal implicit theories work, it seems likely that people vary in their beliefs about the extent to which body weight is something that can be changed and that these beliefs are critical for a host of important outcomes.

Based on the allegiance to weight loss efforts, it seems most people believe weight can change. For example, in the United States, people spend \$40 billion a year on dieting-related products—an amount that is more than the GDP of half of the world's nations (National Eating Disorders Association, 2005). But, what *should* people believe about weight? Should people believe that weight is primarily changeable or should they consider the genetic underpinnings? The answer has implications not only for approaches to combatting obesity, but also influences what people do to manage their weight as well as people's perceptions about others and themselves. The primary goal of the current chapter, in addition to situating the research on weight-based mindsets into the broader work on implicit theories, is to review the empirical evidence illustrating how weight-based mindsets are relevant for understanding individuals' self-regulatory processes and their evaluations of themselves and others.

Weight-Based Mindsets and Self-regulation

Early implicit theory of weight research explored why some people persist and still feel optimistic following inevitable dieting setbacks, whereas others feel helpless and avoid dieting all together. Research suggests that people continue to fail to reach their long-term dietary goals despite early success (Mann et al., 2007). For example, in a meta-analysis of studies examining calorie restricting diets, over one-third of dieters ended up regaining their weight, even after initial successful loss (Mann et al., 2007). However, not all dieters fail, with one study illustrating that approximately 10% or more maintain their weight loss for over a year (Kraschnewski, et al., 2010). Regardless of long-term success rates, one thing is clear—dieters face setbacks. After failing, individuals experience a range of emotional reactions, including assessing whether continuing to strive for the goal is a worthwhile endeavor, and deciding on a course of action. Building on Dweck's (2000) and Ommundsen's (2001) work on the link between implicit theories and feelings of optimism following a setback, early research on implicit theories of weight tested whether feeling more optimistic mediates the link between mindsets and goal striving strategies (Burnette, 2010). More specifically, across three studies,

people with incremental theories of weight demonstrated different goal striving patterns than those with entity theories in the wake of setbacks. In Study 1, in a sample of 250 participants, those with stronger incremental theories of weight reported fewer intentions to avoid future dieting attempts in response to a hypothetical dieting setback and this effect was mediated by more positive expectations about the potential for future success. In Study 2, in a sample of 287 participants currently trying to lose weight, those with stronger incremental theories remained optimistic in the wake of their dieting setbacks and reported less avoidant coping. These processes had downstream implications for weight loss success. In Study 3, an experimental manipulation provided causal evidence of the effects of an incremental theory of weight on expectations and goals striving. Individuals assigned to the *incremental* condition read a paragraph summarizing the evidence that weight can be changed, whereas participants assigned to the *entity* condition read evidence about the genetic basis of weight. Participants in the changeable, relative to the fixed condition, reported feeling more optimistic about the potential for success and greater intentions to engage in effortful regulatory strategies in the wake of a hypothetical setback. Taken together, these studies suggest implicit theories of weight serve as core assumptions that guide cognition and subsequent behavioral intentions (Burnette, 2010).

Building on this work, researchers examined if an intervention designed to promote an incremental theory of weight could be beneficial in helping dieters reach their weight loss goals (Burnette & Finkel, 2012). A fundamental question of this work was to examine if incremental beliefs are an important motivational piece for understanding weight fluctuations in the wake of dieting setbacks. As past implicit theory research has shown, mindsets matter most in times of threats to the self (e.g., failure feedback, setbacks; Burnette et al., 2013). More specifically, the authors compared the incremental theory intervention to a knowledge condition and a true no-treatment control to examine weight loss success. In the incremental theory condition, participants received information about the malleable nature of weight, whereas in the knowledge condition, participants received information about how lifestyle, nutrition, and exercise can influence their health. Similar to other work examining dieting, most participants gained weight over time (Neumark-Sztainer, Wall, Haines, Story, & Eisenberg, 2007; Stice, Cameron, Killen, Hayward, & Taylor, 1999). However, participants in the incremental theory condition, who adopted stronger beliefs in the changeable nature of weight, did not gain weight in the wake of severe dieting setbacks. In conclusion, although the knowledge and incremental theory messages were both effective at buffering against the natural trend toward weight gain, the incremental theory mattered most when dieters reported severe setbacks. The authors note that incorporating incremental beliefs into broader dieting regimens may hold promise for helping individuals maximize their weight loss efforts (Burnette & Finkel, 2012).

In related work, believing obesity is caused by genetics impacts eating behavior, providing additional evidence of the important role mindsets play in predicting goal-directed behaviors. For example, participants randomly assigned to read research articles stressing the genetic nature of weight, compared to those who read

about the influence of social networks, consumed more calories (Dar-Nimrod, Cheung, Ruby, & Heine, 2014). Similarly, messages noting that obesity is a disease, relative to those stressing the changeable nature of body weight, at least indirectly, predicted higher calorie food choices (Hoyt, Burnette, & Auster-Gussman, 2014). Additionally, researchers examined how two theories about the nature of obesity: a diet lay theory (believing calories consumed matter most for weight) and an exercise lay theory (believing calories expended matter most for weight) influence weight-related outcomes (McFerran & Mukhopadhyay, 2013). Across multiple studies, using diverse participants, these implicit theories reliably predicted not only eating behaviors but also actual body mass. More specifically, individuals who held or were induced to believe more strongly in a diet lay theory versus those who held or were induced to believe more strongly in an exercise lay theory, consumed fewer calories and were less likely to be overweight. These studies are interesting because both theories are incremental in nature. That is, one could change their eating habits and/or their exercise behavior. Yet differentiating between these two mindsets is important for understanding eating behaviors and ultimately body weight.

The above studies examining eating behaviors primarily focused on different messages and beliefs about the etiology of obesity (e.g., disease model, social network, exercise, consumption, genetics). However, in other work, researchers specifically examined how believing in the fixed versus changeable nature of weight predicted and influenced eating behaviors. Across two studies, participants with stronger fixed, relative to changeable beliefs about weight consumed more calories from pleasure foods—high-calorie, high-fat foods in lab-based taste tests (Studies 1–2; Ehrlinger, Burnette, Park, Harold, & Orvidas, in press). Additionally, in a third study, participants with fixed, relative to changeable, theories of weight reported less resilient nutrition self-efficacy and this predicted self-reported higher calories consumed from fat. That is, participants who believed weight is fixed also believed that they were less capable of coping with challenging eating situations (e.g., belief that one can stick to healthy eating even in the absence of support, or time to develop the routine; Schwarzer & Renner, 2000) and such self-efficacy doubts undermined healthy eating behavior.

In addition to theories of weight influencing eating behaviors, mindsets about body types have implications for exercise behaviors. For example, women who held more of a growth mindset regarding the malleability of the body, reported exercising more on a weekly and yearly basis in comparison to women holding an entity theory. Interestingly, this relationship between implicit theories of body and exercise behavior was moderated by weight discrepancy (i.e., the difference between self-reported current weight and ideal weight). More specifically, the relation between implicit theories and exercise was stronger for the women who reported a low weight discrepancy. Thus, implicit theories may be especially influential on exercise behaviors for individuals who believe their ideal weight to be rather close to their current weight (Lyons, Kaufman, & Rima, 2013).

Considering the increase in the rate of Americans and individuals worldwide who are overweight or obese (James, Leach, Kalamara, & Shayeghi, 2001), and the

evidence reviewed above of the potential self-regulatory benefits of an incremental theory of weight, continuing to study dieting motivation through cognitive frameworks such as implicit theories can have important implications for public policy messages as well as weight loss interventions. However, one important question to answer first is: What are the implications of these beliefs for obesity stigma? There are potential concerns with encouraging an incremental view of body weight that follows from the stereotype content model (Fiske, Cuddy, Glick, & Xu, 2002) which argues that believing someone is in control of their circumstances (i.e., an incremental theory of weight) can lead to greater prejudice and discrimination (also see Crandall & Martinez, 1996). In addition, what are the implications of these incremental theories of weight for body image concerns and stigma internalization? These are questions that researchers are just starting to answer.

Weight-Based Mindsets and Stigma

Beliefs about the changeable nature of weight can motivate some people as well as political and public policy organizations to shame people to lose weight. For example, the state of Georgia, in the United States, incorporated the following slogans into their strong for life campaigns: “Warning: Big bones didn’t make me this way, big meals did,” and “Warning: Fat prevention begins at home and at the buffet line” (Browner, 2012). Far from the truth, these shaming tactics not only perpetuate stigma, they undermine health (Puhl & Brownell, 2003). The fat acceptance movement (e.g., Health at Every Size) tends to focus on the genetic underpinnings of body weight (e.g., Bacon, 2010). Not only do these messages reflect the way scientists, doctors, and the public might think about the nature of body weight, they also have a significant influence on weight-based bias, including the internalization of such stigma.

Weight-based stigma can have devastating effects on overweight and obese individuals in the form of widespread prejudice and discrimination across domains from employment, to health care, to education (Puhl & Heuer, 2009). Moreover, weight-based prejudice and discrimination can have substantial adverse mental and physical health consequences (Hunger & Major, 2015; Major, Eliezer, & Rieck, 2012; Major, Mendes, & Dovidio, 2013; Puhl & Heuer, 2009, 2010). Weight stigma fits within two of the three primary types of stigmas proposed by Goffman (1963): abominations of the body and blemishes of individual character. That is, it arises from both the belief that overweight and obese individuals have a body that breaches aesthetic norms (Farrell, 2011) and the belief that overweight individuals are lazy and lack self-discipline and are therefore to blame for their excess weight (Carr, Jaffe, & Friedman, 2008; Crandall & Martinez, 1996; Crandall & Schiffhauer, 1998). Although a multitude of factors influence perceptions of overweight and obese individuals, at the root are people’s beliefs about the nature and etiology of weight.

Ample research demonstrates that implicit theories about the fixed versus malleable nature of human attributes can play an important role in prejudice toward those with devalued characteristics. In part, this is driven by how people process information when perceiving others (Erdley & Dweck, 1993; Levy, Stroessner, & Dweck, 1998; Plaks, Stroessner, Dweck, & Sherman, 2001; Poon & Koehler, 2008). Typically, entity theorists have a tendency to process information in terms of specific traits, whereas incremental theorists are more likely to focus on the entire picture and process information related to the dynamics of the situational context (e.g., Bastian & Haslam, 2008). That is, entity theorists are much more likely to engage in lay dispositionalism, the tendency to use traits as the critical unit of analysis in social perception (Chiu, Hong, & Dweck, 1997; Levy, Stroessner, & Dweck, 1998). The implicit theory literature consistently links entity theories to both essentialist thinking and to greater prejudice (e.g., Haslam, Bastian, Bain, & Kashima, 2006; Haslam, Rothschild, & Ernst, 2002; Rydell, Hugenberg, Ray, & Mackie, 2007). By regarding characteristics, such as homosexuality, mental illness, or obesity, as fixed, the stigmatized individual is imbued with an inherent ‘differentness’ that is deemed both serious and persistent, which can promote prejudice (Haslam et al., 2002, 2006; Hegarty & Golden, 2008).

However, research on attribution theory suggests that the link between implicit theories of weight and anti-fat prejudice may not be so simple. Obesity is, in part, a characterological stigma driven largely by blame (DeJong, 1993). As Crandall and Reser (2005) noted, the relationship between viewing overweight people as responsible for their weight and anti-fat prejudice is “one of the best-established relations in the study of attitudes toward fat people” (p. 83, Crandall & Reser, 2005). That is, viewing weight as unchangeable and fixed, as opposed to changeable, can serve to reduce the extent to which people blame overweight individuals for their weight and, in turn, reduce anti-fat prejudice (e.g., Weiner, 1985; Weiner, Perry, & Magnusson, 1988).

Thus, believing that weight stems from invariable, as opposed to flexible, underpinnings has the potential to both *increase* anti-fat prejudice by imbuing overweight individuals with an inherent and undesirable differentness, and to *decrease* prejudice by reducing the extent to which people blame overweight individuals for their weight. That is, weight beliefs can have a double-edged sword effect when it comes to anti-fat prejudice, recently termed the *stigma asymmetry model* (Hoyt, Burnette, Auster-Gussman, Blodorn, & Major, 2016). To test these asymmetric stigma mechanisms driving the relation between fixed beliefs and prejudice (i.e., increase via essentialism, decrease via blame), researchers presented participants with weight-related public health messages either promoting the idea that obese people may not have the ability to lose their excess weight, by labeling obesity a disease, or a message arguing that weight is changeable (Hoyt et al., 2016). Across all three experimental studies, the message that obesity is a disease, relative to a message that weight is changeable, simultaneously strengthened the belief in the unchangeable nature of weight and via this mechanism *increased* anti-fat prejudice, and decreased blame and via this mechanism *decreased* anti-fat prejudice.

These double-edged sword effects associated with believing weight is fixed or malleable extend beyond prejudice against others to how people feel shame over their own body and experience stigma internalization. Recently, researchers have taken a closer look at this by drawing upon the robust literature in attribution theory (Burnette, Hoyt, Auster-Gussman & Dweck, manuscript in preparation). According to Burnette and colleagues, conceptualizing weight as changeable or not influences both beliefs that one is personally accountable for one's weight, as well as beliefs that one can engage successfully in weight regulation. Across two studies, they show that believing weight is changeable both strengthens personal responsibility attributions, thereby *increasing* body shame and self-stigma as well as strengthening the efficacy belief that with effort, one has the capacity to offset the current condition, thereby *decreasing* shame and self-stigma. Additionally, they constructed a public health message designed to eliminate the shame that stems from the responsibility attributions but maintain the benefits of the efficacy beliefs. This message, that stressed *not blaming* people for their current weight combined with the argument that effort and effective strategies can contribute to successful weight regulation, was effective in eliminating the link between incremental theories of weight and increased shame and stigma while maintaining the link between theories and decreased shame and self-stigma through stronger beliefs that one can successfully engage in effortful regulation.

In summary, beliefs about the changeable versus fixed nature of weight are important for understanding both individuals' self-regulatory processes in their pursuit of weight management and the role of weight-based stigma in perceptions and evaluations of others and the self. Specifically, believing more strongly that one's weight and body are changeable helps individuals regulate their expectations in the wake of setbacks, which in turn is related to effortful regulatory strategies. Additionally, believing weight is changeable directly and indirectly predicts more successful weight loss and lower body mass. Such beliefs and messages are also linked to reduced consumption of high-calorie, high-fat foods and increases in exercise behaviors. Moreover, in considering how mindsets contribute to anti-fat prejudice, research shows that stronger incremental theories predict greater blame, and via this mechanism, greater anti-fat attitudes, but they also predict less essentialist thinking and via this mechanism, weaker anti-fat attitudes. A similar double-edged sword effect is demonstrated when examining the link to body shame. Specifically, for body shame, incremental theories predict greater onset responsibility, and via this mechanism, greater shame. They also predict greater offset efficacy, and via this mechanism, less shame. Importantly, Burnette and colleagues (manuscript in preparation) found initial support for a public health message that maximized the benefits of an incremental message but reduced the stigma-related costs. However, before we translate findings into policy, there are quite a few unanswered questions regarding implicit theories of weight, as this is a rather nascent area, with promising areas for future inquiry.

Future Directions and Conclusions

How Are Weight-Based Mindsets Formed?

How do individuals learn and decide whether a human attribute can be changed or not? Past work, within academic achievement contexts, has shown that one way is through praise. For example, students who are told they are smart and talented are more likely to develop fixed mindsets of intelligence, whereas those who are praised for their hard work, for choosing effective strategies, and for learning from others are more likely to develop growth mindsets of intelligence (Yeager & Dweck, 2012). But, how do weight-based mindsets develop? One powerful influence is public health messages related to obesity. However, these messages, as we noted in the introduction, are complicated, nuanced and often incongruent. This is probably because in trying to understand the etiology, it quickly becomes apparent that obesity is one of the most vexing issues of modern life. Merely a half century ago, obesity was rare, then it “spread” to just a few nations before developing throughout the world in just a few decades (Wadden, Brownell, & Foster, 2002). Such a rapid increase begs the question, why? And the potential answers are as varied as the colors of autumn. Is the obesogenic environment—one that encourages the overconsumption of unhealthy foods and limits access to exercise—to blame? Or is the Health at Every Size (HAES) movement’s philosophy more convincing? They claim that 70% of an individual’s weight is dictated by genetics and that unrealistic expectations and social stigma contribute to weight issues (Bacon, 2010). Then, in 2013, the American Medical Association (AMA) labeled obesity a disease—a decision with ramifications for not only how much individuals value health (Hoyt et al., 2014), but also with implications for beliefs about weight (e.g., Dar-Nimrod et al., 2014; Hoyt et al., 2014, 2016).

Despite initial evidence for the implications of various public health messages about the etiology of obesity on weight-based mindsets, more work is still needed and less is known about other potential influences. For example, individual experiences and development likely influence weight-based mindsets. In a systematic review assessing the effectiveness of dietary and exercise interventions, weight regain approached the 50% mark (Curioni & Lourenco, 2005). What are the implications of such fluctuations and disappointments for individuals’ mindsets? On a related note, when do theories of weight develop? Implicit theory research within an academic context suggests that theories of intelligence can be assessed as early as elementary school (Erdley & Dweck, 1993). Additionally, this work highlights how mindsets are especially malleable during times of transitions, with most interventions targeting the transition to middle school (Romero, Master, Paunesku, Dweck, & Gross, 2014) or college (Tamir, John, Srivastava, & Gross, 2007). What transition is key to implement the most powerful weight-based mindset intervention? Is it during the shift from weight loss back to weight gain? Is it early in development? Although interventions and experimental studies have proven effective at shifting weight-based mindsets (Burnette, 2010; Burnette &

Finkel, 2012), most of this work has focused on relatively short time periods. Thus, many questions remain about not only how and when these mindsets develop but also about the durability of weight-based mindset intervention effects.

In addition to public health messages and differences in weight management experiences, media has the potential to influence a host of attitudes, including weight-based mindsets. For example, advertisements presented in Western media portray images of models that, due to photoshop and camera angles, are often unrealistic and in doing so, “today’s media blurs the boundaries between glorified fiction and reality” (Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999, p. 93). Because reaching this thin ideal portrayed in the media is not feasible, body dissatisfaction often ensues (Thompson, et al., 1999). Attempts to change one’s body based on fictitious images will almost inevitably lead to failure, and could contribute to the development of more of a fixed mindset regarding one’s body. Another area for future inquiry is to explore how thin ideals and related media may contribute to beliefs about weight.

In summary, many questions remain about how weight-based mindsets are developed. We have reviewed three influences noting (i) public health messages related to the etiology of obesity, (ii) difficulties with sustained weight loss, and (iii) media portrayals of thin ideals as potential contributing forces. In addition to considering how and when weight-based mindsets develop, another relevant question for future inquiry is whether an incremental message should be encouraged.

Can Effort Be Carried Too Far?

We have shown how an incremental theory promotes greater effortful regulation and stronger weight regulation efficacy, but can effort be carried too far? Might especially strong incremental theories of weight promote excessive optimism? In light of work on false hope (Polivy & Herman, 2002), is it possible that too strong of an incremental theory can lead individuals to suffer from overconfidence? Knowing when to abandon unattainable goals is as important as understanding when to persist. With recent work highlighting the difficulties of long-term weight loss (e.g., Anderson, Konz, Frederich, & Wood, 2001; Wing & Phelan, 2005), is unconditional persistence desirable and might an incremental theory be disadvantageous? Future research should examine this question and more. For example, should an incremental theory of weight be encouraged in children? It seems plausible that encouraging incremental theories of weight, especially in children, can put undue pressure and may shame them. Thus, a more nuanced message is needed—one that reduces blame and responsibility and at the same time does not lead to essentialist thinking or helplessness. Burnette and colleagues (under review) have initial evidence of the benefits of such a message without the costs but more work is needed to understand different avenues for shifting weight-based mindsets and to understand the impact of this message on different audiences. For example,

past work demonstrates that praising effort rather than ability can encourage incremental theories of intelligence (Mueller & Dweck, 1998). In the context of weight management, this likely means encouraging healthy habits, rather than pounds shed. Although theories of weight, and implicit theories more generally, are relatively stable beliefs, a powerful situation or message can highlight one view over the other and influence beliefs at that particular moment and beyond. In summary, future work is needed to empirically investigate how to best send an incremental theory of weight message that has the benefits without the costs and how to do so at different developmental stages and across diverse populations.

Cultural Variations?

Are there cultural similarities in beliefs about weight? In other domains, such as intelligence and morality, people in individualistic and collectivistic cultures hold both entity and incremental theories to similar degrees (Chiu et al., 1997). However, obesity rates and media-driven ideals for thinness vary by culture and likely impact beliefs about the nature of weight management. Might obesity rates themselves predict weight-based mindsets by sending implicit messages (e.g., through exposure to high numbers of obese individuals, weight loss programs) about whether weight can be managed? Moreover, cultural disparities in physical attraction ratings based on weight might influence weight-based mindsets. For instance, past work reveals that there are significant differences in what is considered physically attractive depending on world region. When given the opportunity to choose from a variety of female body shapes, participants from Eastern Europe, Scandinavia, and Western Europe tend to select heavier figures than those from Western Europe and the United States (Swami, et al., 2010). Additionally, exposure to Western media is associated with favoring thinner body ideals (Swami, et al., 2010). Beyond obesity rates and ideals for thinness, other sociocultural influences might influence weight-based mindsets. For example, perceptions of weight and motivation to engage in healthy lifestyle behaviors are largely shaped by societal norms and environmental influences (Strahan, Wilson, Cressman, & Buote, 2006). Sociocultural influences not only shape beliefs about the value of weight loss, they also impact diet and exercise behaviors (Ricciardelli, & McCabe, 2001). Future work should explore additional differences and similarities in beliefs about weight across a range of cultures.

Conclusions

In this chapter, we first briefly reviewed how people's beliefs set up a meaning system within which people operate (Dweck, 2000). We identified a core personal construct, namely implicit theories about whether attributes are fixed or malleable,

and illustrated important consequences of these beliefs across a range of contexts. We focused our analysis on implicit theories of weight, enumerating their implications for self-regulation, perceptions of overweight and obese individuals, and stigma internalization. Research in this area endeavors to answer important questions: What leads some individuals to persist and achieve their weight loss goals and others to feel helpless and avoid dieting and physical exercise all together? What makes people more likely to indulge in pleasure food consumption? What helps people stick to exercise regimens? What types of messages about the nature of obesity decrease stigma and body shame? Results demonstrate that implicit theories of weight (and body) are important for understanding the answers to these questions. More specifically, stronger incremental theories of weight predict and lead to more effortful self-regulatory intentions, can help buffer against weight gain following more severe dieting setbacks, and can help reduce calorie consumption from fat.

Despite these potential self-regulatory benefits, we also illustrated how stronger incremental theories of weight can indirectly both decrease and increase anti-fat attitudes and self-stigma via opposing mechanisms. On the one hand, believing that weight is changeable decreases the extent to which people view weight (and thus overweight and obesity) as entrenched in the very nature of people and promotes the belief that one can engage successfully in weight regulation, thereby decreasing prejudice and shame. However, an incremental belief can also lead people to blame others and themselves for their weight, thus promoting prejudice and shame. Research also suggests that public health messages designed to simultaneously decrease the blame placed on individuals for their weight, while also encouraging a belief in the changeable nature of weight might be most effective in reducing both anti-fat prejudice and shame. Moreover, these messages are likely to foster healthy weight management by both decreasing weight stigma (Major et al., 2012) and maintaining the self-regulatory benefits of an incremental theory of weight (Burnette et al., 2013). As Kelly Brownell, a leader in the field of Food Policy and Obesity once wrote, “It is possible and necessary to fight obesity while showing compassion for people who have it” (Hoffman, Salerno, & Moss, 2012, p. xiii).

In conclusion, research on implicit theories of weight provides empirical evidence illustrating the importance of these beliefs for a range of outcomes, from self-regulation to body shame to weight loss success to weight-based prejudice. Yet, this work is in its infancy with numerous questions for future inquiry. For example, how and when do implicit theories of weight develop and what belief should be emphasized? Are there potential drawbacks to an incremental theory of weight, beyond those associated with blame that must be tempered? Importantly, when it comes to weight management, mindsets matter but they are only one piece to the puzzle. How might the power of a “blame-free” incremental theory of weight be harnessed to maximize critical strategies for weight management, especially those that revolve around managing an obesogenic environment in order to make healthy decisions? We hope the summary provided in this chapter provides an impetus for future work examining these questions about weight-based mindsets.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. doi:10.1016/0749-5978(91)90020-T
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), 888. doi:10.1037/0033-2909.84.5.888
- Anderson, J. W., Konz, E. C., Frederich, R. C., & Wood, C. L. (2001). Long-term weight-loss maintenance: A meta-analysis of US studies. *The American Journal of Clinical Nutrition*, 74(5), 579–584.
- Bacon, L. (2010). *Healthy at every size: The surprising truth about your weight*. Dallas, TX: BenBella Books.
- Bastian, B., & Haslam, N. (2008). Immigration from the perspective of hosts and immigrants: Roles of psychological essentialism and social identity. *Asian Journal of Social Psychology*, 11(2), 127–140. doi:10.1111/j.1467-839X.2008.00250.x
- Blackburn, G. (1995). Effect of degree of weight loss on health benefits. *Obesity Research*, 3(S2), 211s–216s. doi:10.1002/j.1550-8528.1995.tb00466.x
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78, 246–263. doi:10.1111/j.1467-8624.2007.00995.x
- Bråten, I., & Strømsø, H. I. (2005). The relationship between epistemological beliefs, implicit theories of intelligence, and self-regulated learning among Norwegian postsecondary students. *British Journal of Educational Psychology*, 75(4), 539–565. doi:10.1348/000709905X25067
- Browner, E. (2012, February 3). S.F. fat activists oppose Atlanta's anti-obesity campaign. *SF Weekly*. Retrieved from Blogs.sfweekly.com
- Burkley, M., Parker, J., Stermer, S. P., & Burkley, E. (2010). Trait beliefs that make women vulnerable to math disengagement. *Personality and Individual Differences*, 48, 234–238. doi:10.1016/j.paid.2009.09.002
- Burnette, J. L. (2010). Implicit theories of body weight: Entity beliefs can weigh you down. *Personality and Social Psychology Bulletin*, 36, 410–422. doi:10.1177/0146167209359768
- Burnette, J. L., & Finkel, E. J. (2012). Buffering against weight gain following dieting setbacks: An implicit theory intervention. *Journal of Experimental Social Psychology*, 48, 721–725. doi:10.1016/j.jesp.2011.12.020
- Burnette, J. L., Hoyt, C. L., Auster-Gussman, L., & Dweck, C. S. (manuscript in preparation). Believing weight is changeable: Fostering incremental theories of weight that have the benefits without the costs.
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139(3), 655. doi:10.1037/a0029531
- Carr, D., Jaffe, K. J., & Friedman, M. A. (2008). Perceived interpersonal mistreatment among obese Americans: Do race, class, and gender matter? *Obesity*, 16(S2), S60–S68. doi:10.1038/oby.2008.453
- Chen, L. H., Chen, M. Y., Lin, M. S., Kee, Y. H., Kuo C. F., & Shui, S. H. (2008). Implicit theory of athletic ability and self-handicapping in college students. *Psychological Reports*, 103, 476–484. doi:10.2466/pr0.103.2.476-484
- Chiu, C. Y., Hong, Y. Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology*, 73(1), 19. doi:10.1037/0022-3514.73.1.19
- Costa, P. T., & McCrae, R. R. (1985). The NEO personality inventory.
- Crandall, C. S., & Martinez, R. (1996). Culture, ideology, and antifat attitudes. *Personality and Social Psychology Bulletin*, 22(11), 1165–1176. doi:10.1177/01461672962211007
- Crandall, C. S., & Reser, A. H. (2005). Attributions and weight-based prejudice. *Weight bias: Nature, consequences, and remedies* (pp. 83–96). New York, NY: Guilford Press.

- Crandall, C. S., & Schiffhauer, K. L. (1998). Anti-fat prejudice: Beliefs, values, and American culture. *Obesity Research*, 6(6), 458–460. doi:10.1002/j.1550-8528.1998.tb00378.x
- Curioni, C. C., & Lourenco, P. M. (2005). Long-term weight loss after diet and exercise: A systematic review. *International Journal of Obesity*, 29(10), 1168–1174. doi:10.1038/sj.jo.0803015
- Cury, F., Da Fonseca, D., Zahn, I., & Elliot, A. (2008). Implicit theories and IQ test performance: A sequential meditational analysis. *Journal of Experimental Social Psychology*, 44, 783–791. doi:10.1016/j.jesp.2007.07.003
- Dar-Nimrod, I., Cheung, B. Y., Ruby, M. B., & Heine, S. J. (2014). Can merely learning about obesity genes affect eating behavior? *Appetite*, 81, 269–276. doi:10.1016/j.appet.2014.06.109
- DeJong, W. (1993). Obesity as a characterological stigma: The issue of responsibility and judgments of task performance. *Psychological Reports*, 73, 963–970. doi:10.2466/pr0.1993.73.3.963
- Diener, C. I., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36, 451–462. doi:10.1037/0022-3514.36.5.451
- Diener, C. I., & Dweck, C. S. (1980). An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology*, 39(5), 940–952. doi:10.1037/0022-3514.39.5.940
- Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Dweck, C. S. (2012). Implicit theories. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 2, pp. 43–61). Thousand Oaks, CA: Sage.
- Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995a). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, 6(4), 267–285.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995b). Implicit theories: Elaboration and extension of the model. *Psychological Inquiry*, 6, 322–333. doi:10.1207/s15327965pli0604_12
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, 95(2), 256–273. doi:10.1037/0033-295X.95.2.256
- Dweck, C. S., & Molden, D. C. (2005). Self-theories: Their impact on competence motivation and acquisition. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 122–140). New York: Guilford Press.
- Ehrlinger, J., Burnette, J. L., Park, J., Harold, M. L., & Orvidas*, K. (forthcoming). Incremental theories of weight predict lower consumption of high-calorie, high-fat foods. *Journal of Applied Social Psychology*.
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54, 5–12. doi:10.1037/0022-3514.54.1.5
- Erdley, C. A., & Dweck, C. S. (1993). Children's implicit personality theories as predictors of their social judgments. *Child Development*, 64(3), 863–878. doi:10.1111/j.1467-8624.1993.tb02948.x
- Farell, A. E. (2011). *Fat shame: Stigma and the fat body in American culture*. New York, NY: NYU Press.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878. doi:10.1037/0022-3514.82.6.878
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall.
- Haslam, N., Bastian, B., Bain, P., & Kashima, Y. (2006). Psychological essentialism, implicit theories, and intergroup relations. *Group Processes & Intergroup Relations*, 9(1), 63–76. doi:10.1177/1368430206059861
- Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, 41(1), 87–100. doi:10.1348/014466602165072

- Hegarty, P., & Golden, A. M. (2008). Attributional beliefs about the controllability of stigmatized traits: Antecedents or justifications of prejudice? *Journal of Applied Social Psychology, 38*(4), 1023–1044. doi:[10.1111/j.1559-1816.2008.00337.x](https://doi.org/10.1111/j.1559-1816.2008.00337.x)
- Henderson, V., & Dweck, C. (1990). Adolescence and achievement. In S. Feldman & G. Elliott (Eds.), *At the threshold: Adolescent development* (pp. 308–329). Cambridge, MA: Harvard University Press.
- Hoffman, J., Salerno, J. A., & Moss, A. (2012). *The weight of the nation: Surprising lessons about diets, food, and fat from the extraordinary series from HBO documentary films*. London, United Kingdom: Macmillan.
- Hong, Y., Chiu, C., Dweck, C. S., Lin, D., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*, 588–599. doi:[10.1037/0022-3514.77.3.588](https://doi.org/10.1037/0022-3514.77.3.588)
- Hoyt, C. L., Burnette, J. L., & Auster-Gussman, L. (2014). “Obesity is a disease” Examining the self-regulatory impact of this public-health message. *Psychological Science, 25*(4), 997–1002. doi:[10.1177/0956797613516981](https://doi.org/10.1177/0956797613516981)
- Hoyt, C. L., Burnette, J. L., Auster-Gussman, L., Blodorn, A., & Major, B. (2016). The obesity stigma asymmetry model: The indirect and divergent effects of blame and changeability beliefs on antifat prejudice. *Stigma and Health*. Advance online publication. doi:[10.1037/sah0000026](https://doi.org/10.1037/sah0000026)
- Hunger, J. M., & Major, B. (2015). Weight stigma mediates the association between BMI and self-reported health. *Health Psychology, 34*(2), 172–175. doi:[10.1037/hea0000106](https://doi.org/10.1037/hea0000106)
- James, P. T., Leach, R., Kalamara, E., & Shayeghi, M. (2001). The worldwide obesity epidemic. *Obesity Research, 9*(S11), 228S–233S. doi:[10.1038/oby.2001.123](https://doi.org/10.1038/oby.2001.123)
- Kelly, G. A. (1955). *The psychology of personal constructs. Volume 1: A theory of personality*. New York City, NY: WW Norton and Company.
- Kraschnewski, J. L., Boan, J., Esposito, J., Sherwood, N. E., Lehman, E. B., Kephart, D. K., & Sciamanna, C. N. (2010). Long-term weight loss maintenance in the United States. *International Journal of Obesity, 34*(11), 1644–1654. doi:[10.1038/ijo.2010.94](https://doi.org/10.1038/ijo.2010.94)
- Leondari, A., & Gialamas, V. (2002). Implicit theories, goal orientations, and perceived competence: Impact on students’ achievement behavior. *Psychology in the Schools, 39*, 279–291. doi:[10.1002/pits.10035](https://doi.org/10.1002/pits.10035)
- Let’s Move Campaign. (2010). Retrieved from: <http://www.letsmove.gov/>
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology, 74*(6), 1421. doi:[10.1037/0022-3514.74.6.1421](https://doi.org/10.1037/0022-3514.74.6.1421)
- Lyons, C., Kaufman, A. R., & Rima, B. (2013). Implicit theories of the body among college women: Implications for physical activity. *Journal of Health Psychology, 32*, 1142–1153. doi:[10.1177/1359105313508997](https://doi.org/10.1177/1359105313508997)
- Maiman, L. A., Becker, M. H., Kirscht, J. P., Haefner, D. P., & Drachman, R. H. (1977). Scales for measuring health belief model dimensions: a test of predictive value, internal consistency, and relationships among beliefs. *Health Education and Behavior, 5*(3), 215–230. doi:[10.1177/109019817700500303](https://doi.org/10.1177/109019817700500303)
- Major, B., Eliezer, D., & Rieck, H. (2012). The psychological weight of weight stigma. *Social Psychological and Personality Science, 3*(6), 651–658. doi:[10.1177/1948550611434400](https://doi.org/10.1177/1948550611434400)
- Major, B., Mendes, W. B., & Dovidio, J. F. (2013). Intergroup relations and health disparities: A social psychological perspective. *Health Psychology, 32*(5), 514–524. doi:[10.1037/a0030358](https://doi.org/10.1037/a0030358)
- Mann, T. (2015). *Secrets from the eating lab: The science of weight loss, the myth of willpower, and why you should never diet again*. New York City, NY: Harper Collins.
- Mann, T., Tomiyama, A. J., Westling, E., Lew, A. M., Samuels, B., & Chatman, J. (2007). Medicare’s search for effective obesity treatments: Diets are not the answer. *American Psychologist, 62*(3), 220–233. doi:[10.1037/0003-066X.62.3.220](https://doi.org/10.1037/0003-066X.62.3.220)
- Maurer, T. J., Mitchell, D. R. D., & Barbeite, F. G. (2002). Predictors of attitudes toward a 360-degree feedback system and involvement in post-feedback management development activity. *Journal of Occupational and Organizational Psychology, 75*, 87–107. doi:[10.1348/096317902167667](https://doi.org/10.1348/096317902167667)

- McFerran, B., & Mukhopadhyay, A. (2013). Lay theories of obesity predict actual body mass. *Psychological Science*, 0956797612473121. doi:10.1177/0956797612473121
- Molden, D. C., & Dweck, C. S. (2006). Finding meaning in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, 61, 192–203. doi:10.1037/0003-066X.61.3.192
- Moskowitz, G. B., & Grant, H. (2009). *The psychology of goals*. New York: Guilford Press.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33–52. doi:10.1037/0022-3514.75.1.33
- National Eating Disorders Association. (2005). Statistics: Eating disorders and their precursors. Retrieved from: http://www.nationaleatingdisorders.org/p.asp?WebPage_ID=320&Profile_ID=41138
- Neumark-Sztainer, D., Wall, M., Haines, J., Story, M., & Eisenberg, M. E. (2007). Why does dieting predict weight gain in adolescents? Findings from project EAT-II: A 5-year longitudinal study. *Journal of the American Dietetic Association*, 107(3), 448–455. doi:10.1016/j.jada.2006.12.013
- Nichols, J. D., White, J. J., & Price, M. (2006). Beliefs of intelligence, knowledge acquisition, and motivational orientation: A comparative analysis of Hispanic/Latino and Anglo youth. *Multicultural Perspectives*, 8, 39–48. doi:10.1207/s15327892mcp0804_7
- Niiya, Y., Crocker, J., & Bartmess, E. N. (2004). From vulnerability to resilience learning orientations buffer contingent self-esteem from failure. *Psychological Science*, 15(12), 801–805. doi:10.1111/j.0956-7976.2004.00759.x
- Norman, P., Conner, M., & Bell, R. (1999). The theory of planned behavior and smoking cessation. *Health Psychology*, 18(1), 89–94. doi:10.1037/0278-6133.18.1.89
- Ommundsen, Y. (2001). Students' implicit theories of ability in physical education classes: The influence of motivational aspects of the learning environment. *Learning Environments Research*, 4(2), 139–158. doi:10.1023/A:1012495615828
- Ommundsen, Y. (2003). Implicit theories of ability and self-regulation strategies in physical education classes. *Educational Psychology*, 23, 141–157. doi:10.1080/01443410303224
- Pasanisi, F., Contaldo, F., De Simone, G., & Mancini, M. (2001). Benefits of sustained moderate weight loss in obesity. *Nutrition, Metabolism, and Cardiovascular Diseases*, 11(6), 401–406.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, 26(6), 784–793. doi:10.1177/0956797615571017
- Piaget, J., & Garcia, R. (1991). In P. M. Davidson (Ed.) *Toward a logic of meanings*. Berlin, Germany: Erlbaum.
- Plaks, J. E., & Stecher, K. (2007). Unexpected improvement, decline, and stasis: A prediction confidence perspective on achievement success and failure. *Journal of Personality and Social Psychology*, 93(4), 667–684. doi:10.1037/0022-3514.93.4.667
- Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology*, 80(6), 876–893. doi:10.1037/0022-3514.80.6.876
- Polivy, J., & Herman, C. P. (2002). If at first you don't succeed: False hopes of self-change. *American Psychologist*, 57(9), 677–689. doi:10.1037/0003-066X.57.9.677
- Pollack, A. (2013). AMA recognizes obesity as a disease. *The New York Times*, 18.
- Poon, C. S., & Koehler, D. J. (2008). Person theories: Their temporal stability and relation to intertrait inferences. *Personality and Social Psychology Bulletin*, 34(7), 965–977. doi:10.1177/0146167208316690
- Puhl, R. M., & Brownell, K. D. (2003). Psychosocial origins of obesity stigma: Toward changing a powerful and pervasive bias. *Obesity Reviews*, 4(4), 213–227. doi:10.1046/j.1467-789X.2003.00122.x
- Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: A review and update. *Obesity*, 17(5), 941–964. doi:10.1038/oby.2008.636

- Puhl, R. M., & Heuer, C. A. (2010). Obesity stigma: Important considerations for public health. *American Journal of Public Health, 100*(6), 1019–1028. doi:[10.2105/AJPH.2009.159491](https://doi.org/10.2105/AJPH.2009.159491)
- Resnicow, K., Davis-Hearn, M., Smith, M., Baranowski, T., Lin, L. S., Baranowski, J., Doyle, C., & Wang, D. T. (1997). Social-cognitive predictors of fruit and vegetable intake in children. *Health Psychology, 16*(3), 272–276. doi:[10.1037/0278-6133.16.3.272](https://doi.org/10.1037/0278-6133.16.3.272)
- Ricciardelli, L. A., & McCabe, M. P. (2001). Self-esteem and negative affect as moderators of sociocultural influences on body dissatisfaction, strategies to decrease weight, and strategies to increase muscles among adolescent boys and girls. *Sex Roles, 44*(3–4), 189–207. doi:[10.1023/A:1010955120359](https://doi.org/10.1023/A:1010955120359)
- Rippe, J. M., Price, J. M., Hess, S. A., Kline, G., DeMers, K. A., Damitz, S., Kreidieh, I., & Freedson, P. (1998). Improved psychological well-being, quality of life, and health practices in moderately overweight women participating in a 12-week structured weight loss program. *Obesity Research, 6*(3), 208–218. doi:[10.1002/j.1550-8528.1998.tb00339.x](https://doi.org/10.1002/j.1550-8528.1998.tb00339.x)
- Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2014). Academic and emotional functioning in middle school: The role of implicit theories. *Emotion, 14*(2), 227–234. doi:[10.1037/a0035490](https://doi.org/10.1037/a0035490)
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96*(2), 341–357. doi:[10.1037/0033-295X.96.2.341](https://doi.org/10.1037/0033-295X.96.2.341)
- Rydell, R. J., Hugenberg, K., Ray, D., & Mackie, D. M. (2007). Implicit theories about groups and stereotyping the role of group entitativity. *Personality and Social Psychology Bulletin, 33*(4), 549–558. doi:[10.1177/0146167206296956](https://doi.org/10.1177/0146167206296956)
- Schwarzer, R., & Renner, B. (2000). Social-cognitive predictors of health behavior: Action self-efficacy and coping self-efficacy. *Health Psychology, 19*(5), 487–495. doi:[10.1037/0278-6133.19.5.487](https://doi.org/10.1037/0278-6133.19.5.487)
- Spinath, B., Spinath, F. M., Riemann, R., & Angleitner, A. (2003). Implicit theories about personality and intelligence and their relationship to actual personality and intelligence. *Personality and Individual Differences, 35*(4), 939–951. doi:[10.1016/S0191-8869\(02\)00310-0](https://doi.org/10.1016/S0191-8869(02)00310-0)
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629. doi:[10.1037/0003-066X.52.6.613](https://doi.org/10.1037/0003-066X.52.6.613)
- Stice, E., Cameron, R. P., Killen, J. D., Hayward, C., & Taylor, C. B. (1999). Naturalistic weight-reduction efforts prospectively predict growth in relative weight and onset of obesity among female adolescents. *Journal of Consulting and Clinical Psychology, 67*(6), 967–974. doi:[10.1037/0022-006X.67.6.967](https://doi.org/10.1037/0022-006X.67.6.967)
- Strahan, E. J., Wilson, A. E., Cressman, K. E., & Buote, V. M. (2006). Comparing to perfection: How cultural norms for appearance affect social comparisons and self-image. *Body Image, 3*(3), 211–227. doi:[10.1016/j.bodyim.2006.07.004](https://doi.org/10.1016/j.bodyim.2006.07.004)
- Sturm, R., & Hattori, A. (2013). Morbid obesity rates continue to rise rapidly in the United States. *International Journal of Obesity, 37*(6), 889–891. doi:[10.1038/ijo.2012.159](https://doi.org/10.1038/ijo.2012.159)
- Swami, V., Frederick, D. A., Aavik, T., Alcalay, L., Allik, J., Anderson, D., . . . , Danel, D. (2010). The attractive female body weight and female body dissatisfaction in 26 countries across 10 world regions: Results of the International Body Project I. *Personality and Social Psychology Bulletin, 36*(3), 309–325. doi:[10.1177/0146167209359702](https://doi.org/10.1177/0146167209359702)
- Tabernero, C., & Wood, R. E. (1999). Implicit theories versus the social construal of ability in self-regulation and performance on a complex task. *Organizational Behavior and Human Decision Processes, 78*(2), 104–127. doi:[10.1006/obhd.1999.2829](https://doi.org/10.1006/obhd.1999.2829)
- Tamir, M., John, O. P., Srivastava, S., & Gross, J. J. (2007). Implicit theories of emotion: Affective and social outcomes across a major life transition. *Journal of Personality and Social Psychology, 92*(4), 731–744. doi:[10.1037/0022-3514.92.4.731](https://doi.org/10.1037/0022-3514.92.4.731)
- Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (1999). *Exacting beauty: Theory, assessment, and treatment of body image disturbance*. Washington, DC: American Psychological Association.

- Wadden, T. A., Brownell, K. D., & Foster, G. D. (2002). Obesity: Responding to the global epidemic. *Journal of Consulting and Clinical Psychology, 70*(3), 510–525. doi:[10.1037/0022-006X.70.3.510](https://doi.org/10.1037/0022-006X.70.3.510)
- Wang, C. K., & Biddle, S. J. H. (2001). Young people's motivational profiles in physical activity: A cluster analysis. *Journal of Sport & Exercise Psychology, 23*, 1–22.
- Wang, C. K., Chatzisarantis, N. L., Spray, C. M., & Biddle, S. J. (2002). Achievement goal profiles in school physical education: Differences in self-determination, sport ability beliefs, and physical activity. *British Journal of Educational Psychology, 72*(3), 433–445. doi:[10.1348/000709902320634401](https://doi.org/10.1348/000709902320634401)
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573. doi:[10.1037/0033-295X.92.4.548](https://doi.org/10.1037/0033-295X.92.4.548)
- Weiner, B., Perry, R. P., & Magnusson, J. (1988). An attributional analysis of reactions to stigmas. *Journal of Personality and Social Psychology, 55*(5), 738. doi:[10.1037/0022-3514.55.5.738](https://doi.org/10.1037/0022-3514.55.5.738)
- Wilks, D. C., Sharp, S. J., Ekelund, U., Thompson, S. G., Mander, A. P., Turner, R. M., Jebb, S.A. & Lindroos, A. K. (2011). Objectively measured physical activity and fat mass in children: A bias-adjusted meta-analysis of prospective studies. *PLoS One, 6*(2), e17205. doi:[10.1371/journal.pone.0017205](https://doi.org/10.1371/journal.pone.0017205)
- Wing, R. R., & Phelan, S. (2005). Long-term weight loss maintenance. *The American Journal of Clinical Nutrition, 82*(1), 222S–225S.
- Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*(4), 302–314. doi:[10.1080/00461520.2012.722805](https://doi.org/10.1080/00461520.2012.722805)

Lay Theories and Metaphors of Health and Illness

David J. Hauser, Randolph M. Nesse and Norbert Schwarz

How do we understand how our body works? How do we explain what is happening when someone is in pain or is displaying symptoms? How do we infer the best ways to treat disease? We often understand abstract ideas like how the body works and what the disease is by linking them to domains that are easier to understand (Lakoff & Johnson, 1980). These links are referred to as conceptual metaphors, and they are often seen in the metaphors people use to talk about these topics. As we discuss in this chapter, conceptual metaphors are an important source of lay theories of how the body and disease work. For example, enemy metaphors for cancer, such as references to the “War on cancer” or telling others to “fight cancer,” affect how people think about the disease and its prevention. Machine metaphors for the body that assume it is made of discrete parts, each serving a single purpose, guide the focus of medical research. Because metaphors can shape the reasoning of lay persons as well as professionals, medical professionals should be aware of the metaphors they use and the inferences these metaphors invite.

D.J. Hauser (✉)

Department of Psychology, University of Michigan, 3233 East Hall,
530 Church St, Ann Arbor, MI 48109-1043, USA
e-mail: djhauser@umich.edu

R.M. Nesse

Center of Evolution and Medicine, Arizona State University, Tempe, AZ, USA

N. Schwarz

Department of Psychology, Marshall School of Business, Dornsife Mind & Society Center,
University of Southern California, Los Angeles, CA, USA

Conceptual Metaphors and Lay Theories

Because they are not objects that provide direct sensory experience, abstract concepts are inherently difficult to comprehend and represent (Barsalou, 1999; Paivio, 1971, 1986). What do we think of when we think of justice? What picture pops up in our head when we consider morality? Few of us have firm representations of these concepts that automatically activate when we encounter them.

In contrast, most of us have an easier time comprehending and representing concrete concepts, things with which we have direct sensory experience. It is much easier to think of dogs than of justice. A clearer picture pops into mind when we consider apples than morality. Things that we directly encounter are easier to comprehend and have more stable cognitive representations (Barsalou, 1999; Paivio, 1971, 1986).

So then, how do we understand and represent things that we can never directly experience? Conceptual metaphor theory (Lakoff & Johnson, 1980, 1999) suggests that we do so by linking them to easier-to-understand concrete domains. This saves cognitive effort and provides readily available representations of abstract concepts. Rather than struggling to comprehend an abstract concept like time, we instead think about it in terms of a domain we have much more experience with, such as physical space (Boroditsky, 2000; Boroditsky & Ramscar, 2002; Hauser, Carter, & Meier, 2009). We can move future events around in time just like we can move a chair around a room—we can reschedule events and *move them forward or backward* a few days. We can *approach* future events, while we can leave the past *behind us*. These conceptual metaphors bring a host of easy-to-reach inferences about what time is like and how it is structured that make thinking about it much easier. They simplify representations of abstract concepts in terms of easier-to-understand, familiar domains.

Conceptual metaphors are ubiquitous. One can spot them in common conversational conventions. We talk about divinity and valence by drawing upon verticality (Meier, Hauser, Robinson, Friesen, & Schjeldahl, 2007), moral purity by drawing upon cleanliness (Schnall, Benton, & Harvey, 2008; Lee & Schwarz, 2010), friendliness by drawing upon warmth (Williams & Bargh, 2008; IJzerman & Semin, 2009), and importance by drawing upon heaviness (Jostmann, Lakens, & Schubert, 2009; Chandler, Reinhard, & Schwarz, 2012; Hauser & Schwarz, 2015a; see Landau, Meier, & Keifer, 2010, for more examples). Multiple conceptual metaphors can also be used to comprehend the same domain. As alluded to earlier, we use physical space to think about time (Boroditsky, 2000), but we also think about time as if it is money (Lakoff & Johnson, 1980). Time can be *spent*, *stolen*, *wasted*, *exchanged*, *loaned*, etc. This conceptual metaphor makes many of the same ways of thinking about money relevant to thinking about time. Not only can abstract concepts utilize multiple conceptual metaphors, but also each metaphor can explain and highlight different attributes of the abstract concept. Relating time to space highlights the dynamic nature of events within time, such as how they can be moved and how we perpetually draw closer to future events, whereas relating time

to money highlights the value of time, facilitating decisions such as whether I should spend the morning writing a manuscript or watching television. Different conceptual metaphors often provide different explanations of the same abstract concept.

As explained in other chapters in this volume, lay theories are explanations that people use to understand and predict the world. Whether it is a lay theory of intelligence as fixed or malleable (Dweck, 1999), self-control as a depletable resource (Job, 2017), or mind-wandering as controllable or uncontrollable (Zedelius & Schooler, 2017), lay theories provide simplified accounts of complex phenomena. People draw upon these lay theories to understand their own thought processes and the social world, which then guides decisions about their behavior.

Conceptual metaphors have much in common with lay theories. Both provide simplified explanations which people use to inform their reasoning and behavior. In that way, both are adaptive by helping people easily represent ambiguous situations. However, both can be maladaptive; certain lay theories of intelligence and self-control can have negative effects for how we respond to situations (Dweck, 1999; Job, 2017), and as will be shown, conceptual metaphors for cancer have negative effects on how we think about the disease (Hauser & Schwarz, 2015b).

We propose that conceptual metaphors and lay theories are fundamentally related constructs. Most lay theories highlight important attributes of a domain; for example, that intelligence is fixed or malleable (Dweck, 1999) or that mind-wandering is uncontrollable or controllable (Zedelius & Schooler, 2017). Conceptual metaphors operate similarly by highlighting important attributes of the abstract concept that fit the concrete domain (Lakoff & Johnson, 1980, 1999). For example, “time is money” metaphors imply that time is valuable, and as we discuss later, “cancer is an enemy” metaphors imply that attack rather than prevention is important (Hauser & Schwarz, 2015b). Moreover, lay theories often rely upon conceptual metaphors to communicate information. For instance, people holding an incremental view of intelligence often describe it as being a skill that one can grow “like a muscle” (Dweck, 1999), and willpower is often portrayed as a muscle that can get fatigued (Job, 2017). Similarly, conceptual metaphors can provide explanations of abstract concepts (i.e., lay theories) that people use to understand and predict the world. For example, conceptual metaphors of valence imply that positivity is up and negativity is down (Meier & Robinson, 2004), which creates preferences for Northern real estate locations (Meier, Moller, Chen, & Reimer-Peltz, 2011) and more vote choice for higher listed politicians in election ballots (Kim, Krosnick, & Casasanto, 2015). Conceptual metaphors tend to represent general concepts (e.g., the body, time, and valence) and are applicable across many different domains while lay theories tend to explain more specific concepts (e.g., intelligence and willpower). However, both function as accessible concepts that influence comprehension, representation, and inference when they are applicable to a target. In sum, conceptual metaphors and lay theories have similar effects, in that they both suggest important qualities of an abstract domain, which carries implications for how people interpret ambiguous situations and ultimately behave.

These functional parallels do not imply that conceptual metaphors and lay theories are fundamentally indistinguishable from each other. Each arises through different processes, and the research traditions surrounding each rarely intersect. But when it comes to how people think about health and disease, conceptual metaphors and lay theories operate similarly. They each suggest a crucial attribute about a phenomenon that can guide reasoning and behavior.

Conceptual Metaphors of Health and Disease

Bellicose Metaphors and Cancer

America is *waging a war* on cancer. Many patients are *fighting* and *battling* the disease. Cancerous cells *attack* our bodies, and we try to *destroy* them with treatment. Former President Obama claimed that “now is the time to commit ourselves to *waging a war* against cancer *as aggressive as the war that cancer wages against us*” (emphasis added; Lennon, 2009).

As shown here, public discourse commonly relates cancer to a hostile enemy invader that we must attack and defeat (Sontag, 1978). It is the most popular conceptual metaphor employed by science journalists (Camus, 2009) and cancer patients in online forums (Semino et al., 2015). In everyday American discourse, two of the top ten verbs are used to describe cancer call upon this metaphor (*fight* cancer and *battle* cancer; Davies, 2008, cited in Hauser & Schwarz, 2015b).

However, these bellicose metaphors were not always dominant. They surfaced only in the 1930s, when Mary Lasker joined the American Cancer Society and used the metaphor to lobby for research funding. Her efforts were ultimately successful, as the National Cancer Act of 1971 was passed, marking the beginning of the government’s “War on Cancer” (Mukherjee, 2010). Thus, war metaphors began as a way to drum up support for cancer research, and they have been effective.

Somewhere along the way, however, these metaphors transformed how people think about cancer in general. Many metaphors center on how the public should act toward cancer, such as sayings like, “We need to *fight* and *beat* cancer.” Other sayings portray cancer biology and treatment in bellicose terms, positing that cancerous cells *attack* the body and that our cancer treatments are the *bullets* and *ammunition* that we use to *destroy* cancer. Others link cancer prevention to the metaphor, like books that detail the foods that *fight* cancer or the cancer *fighting* diet (Beliveau & Gingras, 2006). These bellicose expressions are interesting examples of how metaphors may become overextended. While they had positive effects on support for cancer research, they have now come to explain other aspects of cancer, such as its biology, treatment, and prevention, with potential negative effects (Aktipis, Maley, & Neuberg, 2010).

Simple exposure to this pervasive metaphoric language may negatively affect how people think about cancer. Disease is an abstract concept. While people often

experience the symptoms of illness in a very concrete way (e.g., a runny nose, a sore throat, a high fever), this is not the case for the underlying *cause* of such symptoms, the disease, which remains an abstract and ambiguous source of pain. Encountering a metaphorically framed abstract concept induces people to think about it in terms of the metaphor (Lakoff & Johnson, 1980; Landau, Sullivan, & Greenberg, 2009; Gibbs, 2014; Ottati, Renstrom, & Price, 2014; Thibodeau & Boroditsky, 2011). This highlights attributes of the concept that fit with the source domain and de-emphasizes attributes that are not relevant to the source domain. For instance, reading that crime is a *beast* that *attacks* a city leads people to propose more punitive solutions to a crime wave while reading that crime is a *virus* that *infects* a city leads people to propose more reformative solutions (Thibodeau & Boroditsky, 2011). Thus, exposure to bellicose metaphoric framings of cancer may lead people to think of cancer as an enemy, prompting them to map their knowledge of enemies (such as how they arise, how to deal with them, etc.) onto how they think about cancer. Dealing with enemies tends to encourage active behaviors, traditionally thought of as masculine, aimed at attacking and defeating the enemy. In contrast, limitation and restraint are de-emphasized as ways to deal with enemies. One does not often see limitation and restraint nominated as effective ways to attack enemies (see pilot collocation study in Hauser & Schwarz, 2015b for more information). So, exposure to bellicose metaphors should de-emphasize limitation and restraint as effective ways to deal with cancer.

This metaphoric representation of cancer is only beneficial if it closely maps onto the actual effective ways to deal with cancer. Unfortunately, it does not. Many effective cancer prevention behaviors involve limiting behaviors that are known to increase the risk of cancer, such as smoking, sunbathing, and eating foods associated with cancer. Processed foods, red meats, high fat foods, and high calorie foods increase the risk of cancer, and limiting their intake decreases the risk (World Cancer Research Fund & the American Institute for Cancer Research, 2007; Kushi et al., 2012). Viewing cancer as war against an invader may decrease motivation for these effective prevention behaviors. Because it does not make sense to limit yourself in order to fight enemies, it might not make sense to limit yourself in order to fight cancer. These bellicose framings may convey beliefs about prevention that is ultimately harmful for public health.

Does simply reading bellicose metaphors affect how people think about cancer? To test this possibility, we (Hauser & Schwarz, 2015b, Study 1) recruited 64 participants for an online survey about cancer. We gave them information to read about cancer that was manipulated to either use surface metaphoric utterances relating cancer to a *hostile enemy* that needs to be *fought* or not (randomly assigned). For example, for our enemy metaphors group, participants read that “Cancer is a broad group of disease characterized by the *hostile* growth and *invasive* spread of abnormal cells,” whereas for our control group, the italicized words were eliminated. Then participants were asked what cancer prevention behaviors they could think of. This question was also metaphorically framed. For participants in the enemy metaphors group, the question ended by asking “what things would you do to fight against developing cancer” while for the control group, the question

ended by asking “what things would you do to reduce your risk of developing cancer.” Participants listed behaviors on separate lines and coders rated whether those behaviors involved limiting a behavior known to increase the risk of cancer or increasing a behavior known to decrease the risk of cancer.

As expected, enemy metaphors lessened how often people thought about limiting risk-increasing behaviors but did not affect how often people thought about risk-decreasing behaviors. Participants in the control condition listed, on average, two risky behaviors to limit while participants who read that cancer was *hostile* and *invasive* listed, on average, only one and a half risky behaviors to limit, a significant 25% reduction. As hypothesized, reading bellicose metaphors leads people to bring attributes of enemies to bear on their representation of cancer. Because it does not make sense to limit yourself in order to fight enemies, limiting yourself does not come to mind as a way to fight cancer. These metaphors portraying cancer as a battle decrease cognitive access to effective prevention behaviors.

Do these bellicose metaphors only affect how people think about cancer, or do they also affect what people intend to do? We addressed this question in a second study (Hauser & Schwarz, 2015b, Study 2) on 300 people who took an online survey on health messages. Participants were randomly assigned to read one of three similar messages about colorectal cancer, modeled upon information disseminated by the American Cancer Society (*Cancer Facts & Figures*, American Cancer Society, 2012). The control message used few metaphors and discussed colorectal cancer and its risk statistics. The enemy metaphor message added only six words to the control message, and those words framed cancer as a *hostile enemy*. Finally, the imbalance metaphor message added only five words to the control message, and those words metaphorically framed cancer as *imbalance*.

After reading one of those three messages (randomly assigned), participants rated how much they intended to limit behaviors known to increase their risk of colorectal cancer (limit eating red meats, limit eating high fat foods, etc.). Consistent with the results of Study 1, bellicose metaphors lessened people’s intention to limit risky behaviors. Participants who read a message saying that cancer was *hostile* had less intention to limit risk-increasing behaviors compared to participants who read the control message and participants who read the imbalance metaphor message. Thus, the effect is not a metaphoric framing effect in general, but rather it is specific to talking about cancer as if it is a *hostile enemy* that needs to be *fought*. Exposure to bellicose metaphors not only undermines how often limitation comes to mind, but it also undermines people’s intentions to limit risk-increasing behaviors.

These studies point to the power that conceptual metaphors have over people’s beliefs of how cancer works, and a metaphoric fit mechanism seems to drive the effects. Statements that draw upon the same metaphors are processed more fluently than statements that draw upon metaphors from different domains. When the attributes of a concept clash with someone’s metaphoric representation, processing is hindered (Thibodeau & Durgin, 2008, 2011). Therefore, an enemy representation of cancer only advocates ways to deal with cancer that have the same attributes as ways to fight an enemy. Because limitation by default does not seem like an

effective way to fight an enemy, an enemy representation undermines accessibility of it and intentions for it.

A third study (Hauser & Schwarz, 2015b, Study 3) provided further evidence of this process, demonstrating that the negative effect of bellicose metaphors is eliminated when limitation strategies are framed in a way that fits with fighting enemies. When risk-increasing behaviors were framed as “weakening the body’s ability to fight cancer,” then the negative effect of the enemy metaphor message was eliminated. Thus, limitation is not often thought of as a way to fight enemies, but if one frames it as such, then it motivates people to lessen behaviors that are said to “weaken one’s ability to fight.”

Bellicose metaphors for cancer serve as an interesting example of how metaphors can go awry. They began as a way to motivate funding for cancer research, and they were quite successful at that (Mukherjee, 2010). However, the metaphors were extended into how people think about cancer biology, treatment, and prevention, where they imply that many effective prevention behaviors for cancer are ineffective. Simple exposure to these metaphors undermines the extent to which people think of several prevention behaviors and undermines whether people intend to engage in them. Everyday language has the power to shape people’s beliefs of how cancer works and affects what people intend to do about it.

Machine Metaphors and the Body

The metaphor of the body as a machine looms large in the study of human biology and modern medicine (for reviews, see Osherson & Amarasingham, 1981; Nesse, 2016). It originated in the renaissance but became dominant during the industrial revolution of the eighteenth century, when empirical observation and mechanistic causal principles displaced vitalism (Westfall, 1977). Discoveries from medical dissection allowed philosophers such as Rene Descartes to draw parallels between the body and mechanical contrivances (Cottingham, Stoothoff, & Murdoch, 1984). Blood vessels appeared to operate similarly to hydraulic tubes. Muscles and bones functioned similarly to pulleys and levers. The body was explained in reductionist terms, with each body part serving a specific function and different body parts interacting with one another to form a functioning machine (Miller, 1978).

Medical discourse still relies on this metaphor today (Ochsner, 2010; Mumford, 1974), and it is seen often in how people talk about their ailments. Many diseases or ailments are described as due to broken or malfunctioning parts of the body. People often say they have *bad* shoulders or *bad* knees to describe chronic pain. Type I diabetes is said to stem from a *malfunctioning* pancreas, and cancerous tumors are even sometimes explained as resulting from *malfunctioning* cells which are *stuck* in “divide” mode.

This conceptual metaphor guides people’s beliefs, leading them to bring attributes of machinery to bear upon how the body works (Lakoff & Johnson, 1980, 1999). It implies that, like machines, the body is comprised of discrete parts,

and each part has a specific purpose. These parts interact with one another in order to produce our health, as well as our thoughts and behavior. Disease and ailments are caused by malfunctioning, broken, or worn out parts (Osherson & Amarasingham, 1981). However, many aspects of how the body works do not fit with lay beliefs provided by machinery metaphors (Nesse & Williams, 1994). To that end, these metaphors oversimplify how the body works (Nesse, 2016).

For instance, the different parts of a machine often serve one or a few specific purposes, so these metaphors also suggest that each body part has only a few specific purposes. Again, this is an oversimplification. For instance, the stress system exists not only to mediate arousal in dangerous situations, it also adjusts the body to cope in any situation when activity is necessary (Nesse, Bhatnagar, & Ellis, 2016). Also, some of its actions are not direct but instead function to limit damage that would otherwise be imposed by other aspects of the stress system (Munck, Guyre, & Holbrook, 1984). Even the eyebrows serve multiple functions, including nonverbal signaling and keeping sweat from the eyes.

Machines are designed by engineers who conceived of specific parts with specific functions. They also build in redundancies (i.e., backup systems) in case one part fails. In contrast, natural selection shapes systems with multiple, intimately interacting parts that allow the whole system to continue functioning even if one aspect fails. Instead of specific backup systems, bodies have networks of systems (Kitano, 2004). This has clinical implications. For instance, fever is useful to counter infection, but using drugs to block it is often safe because fever is part of a network of other defenses, including immune responses, cough, etc. Thus, demonstrating that drugs that block fever are safe does not imply that fever is an unnecessary epiphenomenon.

The main shortcoming of the machinery conceptual metaphor arises from evolution. Machines are designed, meaning that parts and processes that were no longer efficient or useful would be removed and replaced with new ones. Humans, on the other hand, are evolved (Darwin, 1871). Natural selection is limited to tinkering in ways that leave bodies with many vulnerabilities, such as the opening of the windpipe into the pharynx, where it can be obstructed by food, and the extraordinarily roundabout path taken by the laryngeal nerve down into the thorax before it ascends to the vocal cords, making it prone to damage by thyroid injury or thyroid surgery. Similarly, the spine evolved to serve four-legged creatures. Standing upright must have given hugely adaptive advantages given the manifold of problems it encourages, including back pain, hemorrhoids, hernias, varicose veins, and knee and ankle pain (Clevenger, 1884; Pennisi, 2012). Natural selection works to reduce such problems, but only on a scale of millions of years, and subject to severe constraints such as the inability to start a new design from scratch. Much modern human disease results from the marvelous environments we have created to meet our every need. For instance, our appetite regulation mechanisms were shaped in environments where fat, salt, and sugar were in short supply, so the obesity epidemic is now overwhelming (Pijl, 2011).

The machine metaphor encourages reductionism and an exclusive focus on the body's mechanisms, resulting in extraordinary overinvestments in research at the

cellular level, and underinvestment in studies of environmental factors. For instance, the vast bulk of research on multiple sclerosis focuses upon brain mechanisms, but rates are an order of magnitude higher in modern environments (Correale & Farez, 2007). Explaining these differences in prevalence should be a major priority. Human populations that routinely have high burdens of worms have very low rates of multiple sclerosis (Correale & Farez, 2007), but the exact nature of the relationship remains unclear even as most research continues to focus on biochemical and immunological mechanisms.

The machine metaphor has also encouraged a model of the brain that has not lived up to empirical observation. For decades, researchers believed that the brain operated similarly to a machine; each part of the brain executes a specific function. Many also believed that specific areas of the brain would map on to specific concepts, with some parts “activating” during thought of a single concept (see discussion of localization in Kosslyn & Andersen, 1995). For instance, the amygdala has been thought to be devoted to fear learning (LeDoux, 2003); however, its functions turn out to be far more diverse, being involved in social learning, self-control, aggression, and reward learning (Balleine & Killcross, 2006). Research proposed a brain location for processing faces (the fusiform face area; Kanwisher, McDermott, & Chun, 1997), but the same area also is activated when car experts look at pictures of cars and when bird watchers look at pictures of birds (Gauthier, Skudlarski, Gore, & Anderson, 2000). While the machine metaphor encouraged models of a localized brain, it appears that brain regions serve multiple functions and are often neither necessary nor sufficient for any one specific function.

The idea that the body is a machine has led to many improvements in medicine, such as a departure from animism. However, it fosters beliefs about the body that oversimplify ideas of how it works and how disease works. Because humans are evolved, they are fundamentally different from well-designed machines (Nesse & Williams, 1994; Nesse, 2016). These beliefs about the body guide medical research in potentially problematic ways that can potentially be prevented by recognizing that the organic complexity of the body is fundamentally different from the designed complexity of machines (Nesse, Ganten, Gregory, & Omenn, 2012).

Awareness of Metaphors in Practice

Conceptual metaphors provide beliefs about complex issues and concepts (Lakoff & Johnson, 1980). They are useful, but their costs can be considerable. Enemy metaphors for cancer suggest that prevention strategies based on limitation are unimportant (Hauser & Schwarz, 2015b). Machine metaphors for the body suggest the body’s mechanisms are like those designed by an engineer, when in fact they are characterized by fundamentally different organic complexity (Osherson & Amarasingham, 1981; Nesse & Williams, 1994; Nesse, 2012, 2016). These beliefs guide how people behave. Enemy metaphors spur research investment and active efforts to conquer cancer; however, they also have negative effects on cancer

prevention. Machine metaphors have been essential to escape vitalism and motivate research to delve more deeply into reductionist detail; however, they misrepresent the organic complexity of the body in fundamental ways.

Human minds rely on metaphors, so there is no escaping them (Lakoff & Johnson, 1980, 1999). However, medical professionals and organizations should be more cognizant of the metaphors they use. Increasing our recognition of their applications in medicine and disease may help to maximize their benefits and minimize their costs (Hauser & Schwarz, 2016). In particular, even though these metaphors may be intended as simple linguistic flourishes, casual use of metaphors by medical professionals can result in widespread misimpressions about diseases.

More attention to the use of metaphors in medicine also provides opportunities to offer superior metaphors. For instance, the long-standing belief that bacteria are bad invaders is being replaced quickly by recognition that they are beneficial partners in an ecosystem that is required for our guts, and our bodies more generally, to function normally (McFarland, 2006). Also, while negative emotions continue to be viewed as abnormal, recognizing that they are states shaped by natural selection because of their benefits offers opportunities to study them with greater sophistication in ways that hopefully will reduce the suffering and stigma of anxiety disorders and depression.

Preferably, medical professionals, educators, and even popular media should try to avoid metaphors that promote harmful inferences and use only ones that promote helpful or more accurate inferences about the body. It is not safe to assume that metaphors that worked in one domain will also work in another. This was the issue with enemy metaphors for cancer; these metaphors worked well for securing funding for cancer research (Mukherjee, 2010), but when applied to the biology of cancer genesis and how to prevent and treat cancer, they promote harmful inferences about cancer prevention (Hauser & Schwarz, 2015b).

Because the effects of conceptual metaphors are multifaceted, investigating the benefits and drawbacks of conceptual metaphors in medicine should prove to be a rich area for future research to explore. For instance, while we documented drawbacks of bellicose metaphors for cancer, similar metaphors have been shown to increase people's willingness to get vaccinated for the flu (Scherer, Scherer, & Fagerlin, 2014). Additionally, first-person shooter games where players virtually "battle" cancerous cells have been shown to increase young patients' adherence to treatment and sense of self-efficacy (Kato, Cole, Bradlyn & Pollock, 2008). On the other hand, war metaphors can have negative implications for patients (for a review, see Hauser & Wassersug, 2015). Many patients simply do not see the metaphor as an apt description of their situation and instead chose to describe it using different metaphors (Reisfield & Wilson, 2004). While many cancer patients "lose the battle," this may also reinforce the idea that they did not "fight hard enough" as the metaphor implies. Such metaphors may also encourage overly aggressive treatments (Aktipis et al., 2010) when palliative care may be a better option for increasing quality of life and extending longevity. There are also a variety of metaphors that people use to understand the body and disease. For instance, the idea that the body is a temple (Synnott, 1992) or that disease is caused by environmental

poisons (Lakoff & Johnson, 1999) may promote more holistic views of the body that emphasize preventative measures rather than reactive treatment-focused ones.

Given the pervasiveness of metaphor in shaping how we think about the body and disease, and its implications, positive and negative, for medical research and treatment, intensive research on the role of metaphor in medicine is needed. Research should focus on the inferences that people draw from metaphorical language in order to assure that metaphorical language does not undermine public health or research efforts. Research can elucidate the metaphors that promote healthy behaviors or ones that closely map onto important or accurate features of a domain. Shedding light on the optimal (and suboptimal) metaphors for the body may provide a rich agenda for future collaborations between scholars in medicine, psychology, and communication in the coming years (Hauser & Schwarz, 2016).

Conclusion

Conceptual metaphors provide lay theories of how the body and how disease work. However, they can ultimately oversimplify these abstract, complex domains. Beliefs that cancer is an enemy make certain prevention behaviors seem less effective. Theories that the body is a machine spur oversimplified beliefs about how the body works. Thus, professionals should be aware of the metaphors they use, and research should investigate what these metaphors imply and how they affect health-related decisions and behaviors.

References

- Aktipis, C. A., Maley, C. C., & Neuberg, S. L. (2010). Psychological barriers to evolutionary thinking in medicine. *Evolution and Medicine Review*. Retrieved from <https://evmedreview.com/psychological-barriers-to-evolutionary-thinking-in-medicine/>
- American Cancer Society. (2012). *Cancer facts & figures 2012*. Atlanta, GA: American Cancer Society.
- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, 22, 637–660.
- Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: An integrated view of amygdala function. *Trends in Neurosciences*, 29(5), 272–279.
- Beliveau, R., & Gingras, D. (2006). *Foods that fight cancer: Preventing cancer through diet*. Toronto, Ontario, Canada: McClelland & Stewart.
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75(1), 1–28.
- Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, 13(2), 185–189.
- Camus, J. T. W. (2009). Metaphors of cancer in scientific popularization articles in the British press. *Discourse Studies*, 11, 465–495.

- Chandler, J. J., Reinhard, D., & Schwarz, N. (2012). To judge a book by its weight you need to know its content: Knowledge moderates the use of embodied cues. *Journal of Experimental Social Psychology*, *48*(4), 948–952.
- Clevenger, S. V. (1884). Disadvantages of the upright position. *The American Naturalist*, *18*(1), 1–9.
- Correale, J., & Farez, M. (2007). Association between parasite infection and immune responses in multiple sclerosis. *Annals of Neurology*, *61*(2), 97–108.
- Cottingham, J., Stoothoff, R., Murdoch, D. (1984). *The philosophical writings of Descartes*, 2 vols. Cambridge, England: Cambridge University Press.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. London, UK: John Murray.
- Davies, M. (2008). *The corpus of contemporary American English: 450 million words, 1990-present*. Retrieved from <http://corpus.byu.edu/cocaf>
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
- 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*(2), 191–197.
- Gibbs R. W. Jr. (2014). Conceptual metaphor in thought and social action. In M. J. Landau, M. D. Robinson, & B. P. (Eds.), *The power of metaphor: Examining its influence on social life* (pp. 17–40). Washington, DC: APA.
- Hauser, D. J., Carter, M. S., & Meier, B. P. (2009). Mellow Monday and furious Friday: The approach-related link between anger and time representation. *Cognition and Emotion*, *23*(6), 1166–1180.
- Hauser, D. J., & Schwarz, N. (2015a). Elaborative thinking increases the impact of physical weight on importance judgments. *Social Cognition*, *33*(2), 120.
- Hauser, D. J., & Schwarz, N. (2015b). The war on prevention: Bellicose cancer metaphors hurt (some) prevention intentions. *Personality and Social Psychology Bulletin*, *41*(1), 66–77.
- Hauser, D. J., & Schwarz, N. (2016). Medical metaphors matter: Experiments can determine the impact of metaphors on bioethical issues. *American Journal of Bioethics*, *16*, 18–19.
- Hauser, D., & Wassersug, R. (2015, March 22). Do we need to end the ‘war’ on cancer? Retrieved from <https://www.theguardian.com/commentisfree/2015/mar/22/saying-youre-fighting-a-war-on-cancer-could-make-you-lose-it>
- IJzerman, H., & Semin, G. R. (2009). The thermometer of social relations mapping social proximity on temperature. *Psychological Science*, *20*(10), 1214–1220.
- Job, V. (2017). Effects of lay theories on self-control. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.
- Jostmann, N. B., Lakens, D., & Schubert, T. W. (2009). Weight as an embodiment of importance. *Psychological Science*, *20*(9), 1169–1174.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.
- Kato, P. M., Cole, S. W., Bradlyn, A. S., & Pollock, B. H. (2008). A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. *Pediatrics*, *122*(2), e305–e317.
- Kim, N., Krosnick, J., & Casasanto, D. (2015). Moderators of candidate name-order effects in elections: An experiment. *Political Psychology*, *36*, 525–542.
- Kitano, H. (2004). Biological robustness. *Nature Reviews Genetics*, *5*(11), 826–837.
- Kosslyn, S. M., & Andersen, R. A. (1995). *Frontiers in Cognitive Neuroscience*. Cambridge, MA: MIT Press.
- Kushi L. H., Doyle C., McCullough M., Rock C. L., Demark-Wahnefried W., Bandera E. V., et al. (2012). American Cancer Society guidelines on nutrition and physical activity for cancer prevention. *CA: A Cancer Journal for Clinicians*, *62*, 30–67.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: University of Chicago Press.

- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic Books.
- Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin*, *136*(6), 1045.
- Landau, M. J., Sullivan, D., & Greenberg, J. (2009). Evidence that self-relevant motives and metaphoric framing interact to influence political and social attitudes. *Psychological Science*, *20*, 1421–1427.
- LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, *23*(4–5), 727–738.
- Lee, S. W., & Schwarz, N. (2010). Dirty hands and dirty mouths: Embodiment of the moral-purity metaphor is specific to the motor modality involved in moral transgression. *Psychological Science*, *21*(10), 1423–1425.
- Lennon C. (2009, June 3). Ovarian cancer: Fighting for a cure. Harper's Bazaar. Retrieved from <http://www.harperbazaar.com/beauty/health-wellness-articles/barack-obama-ovarian-cancer>
- McFarland, L. V. (2006). Meta-analysis of probiotics for the prevention of antibiotic associated diarrhea and the treatment of *Clostridium difficile* disease. *The American Journal of Gastroenterology*, *101*(4), 812–822.
- Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What's "up" with God? Vertical space as a representation of the divine. *Journal of Personality and Social Psychology*, *93*(5), 699–710.
- Meier, B. P., Moller, A. C., Chen, J. J., & Riemer-Peltz, M. (2011). Spatial metaphor and real estate: North-South location biases housing preference. *Social Psychological and Personality Science*, *2*(5), 547–553.
- Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up associations between affect and vertical position. *Psychological Science*, *15*(4), 243–247.
- Miller, J. (1978). *The body in question*. New York, NY: Random House.
- Mukherjee, S. (2010). *The emperor of all Maladies: A biography of cancer*. New York, NY: Simon & Schuster.
- Mumford, L. (1974). *The myth of the machine [Vol. 2], The pentagon of power*. New York, NY: Harcourt, Brace & World.
- Munck, A., Guyre, P. M., & Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Reviews*, *5*, 25–44.
- Nesse, R. M. (2012). *Natural selection is simple, but the systems it shapes are unimaginably complex [Web log post]*. Retrieved from <https://www.edge.org/response-detail/10460>
- Nesse, R. M. (2016, February 4). *The body is not a machine [Web log post]*. Retrieved from <https://evmed.asu.edu/blog/body-not-machine>
- Nesse, R. M., Bhatnagar, S., & Ellis, B. (2016). Evolutionary origins and functions of the stress response system. In G. Fink (Ed.), *Stress: Concepts, cognition, emotion, and behavior: Handbook of stress series* (Vol. 1, pp. 95–101). London, UK: Academic Press.
- Nesse, R. M., Ganten, D., Gregory, T. R., & Omenn, G. S. (2012). Evolutionary foundations for molecular medicine. *Journal of Molecular Medicine (Berlin, Germany)*, *90*(5), 509–522.
- Nesse, R. M., & Williams, G. C. (1994). *Why we get sick: The new science of Darwinian Medicine*. New York, NY: Random House.
- Ochsner, A. (2010). Thoughts on the human body. *The Ochsner Journal*, *10*(1), 44–47.
- Osherson, S., & AmaraSingham, L. (1981). The machine metaphor in medicine. In E. G. Mishler, L. R. Amarasingham, S. D. Osherson, S. T. Hauser, N. E. Waxler, & R. Liem (Eds.), *Social contexts of health, illness, and patient care* (pp. 218–249). Cambridge, England: Cambridge University Press.
- Ottati V., Renstrom R. A., Price E. (2014). The metaphoric framing model: Political communication and public opinion. In M. J. Landau, M. D. Robinson, & B. P. Meier (Eds.), *The power of metaphor: Examining its influence on social life* (pp. 179–202). Washington, DC: APA.
- Paivio, A. (1971). *Imagery and verbal processes*. New York, NY: Holt, Rinehart, & Winston.

- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.
- Pennisi, E. (2012). The burdens of being a biped. *Science*, 336(6084), 974.
- Pijl, H. (2011). Obesity: Evolution of a symptom of affluence. *Netherlands Journal of Medicine*, 69(4), 159–166.
- Reisfield, G. M., & Wilson, G. R. (2004). Use of metaphor in the discourse on cancer. *Journal of Clinical Oncology*, 22(19), 4024–4027.
- Scherer, A. M., Scherer, L. D., & Fagerlin, A. (2014). Getting ahead of illness: Using metaphors to influence medical decision making. *Medical Decision Making*, 35, 37–45.
- Schnall, S., Benton, J., & Harvey, S. (2008). With a clean conscience cleanliness reduces the severity of moral judgments. *Psychological Science*, 19(12), 1219–1222.
- Semino, E., Demjén, Z., Demmen, J., Koller, V., Payne, S., Hardie, A., et al. (2015). The online use of violence and journey metaphors by patients with cancer, as compared with health professionals: A mixed methods study. *BMJ Supportive & Palliative Care*. doi:10.1136/bmjspcare-2014-000785
- Sontag, S. (1978). *Illness as metaphor*. New York, NY: Farrar, Straus & Giroux.
- Synnot, A. (1992). Tomb, temple, machine and self: The social construction of the body. *British Journal of Sociology*, 43, 79–110.
- Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE*, 6, e16782.
- Thibodeau, P. H., & Durgin, F. H. (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. *Journal of Memory and Language*, 58, 521–540.
- Thibodeau, P. H., & Durgin, F. H. (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol*, 26, 206–226.
- Westfall, R. S. (1977). *The construction of modern science: Mechanisms and mechanics*. Cambridge, England: Cambridge University Press.
- Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322(5901), 606–607.
- World Cancer Research Fund & American Institute for Cancer Research. (2007). *Food, nutrition, physical activity, and the prevention of cancer: A global perspective*. Washington, DC: American Institute for Cancer Research.
- Zedelius, C. M., & Schooler, J. W. (2017). Lay theories about spontaneous stimulus-unrelated thoughts. In C. M. Zedelius, B. C. N. Müller, & J. W. Schooler (Eds.), *The science of lay theories—How beliefs shape our cognition, behavior, and health*. New York, NY: Springer.

How Lay Theories Influence Our Mental Health

Adrian Furnham

This chapter reviews the scattered research on what people think are the causes, manifestations, consequences, and cures for general and specific mental health problems. People come to develop ‘theories’ about their own and others’ mental and physical health from various sources including the media, formal and informal education, as well as personal experiences of illness and health. On the basis of these theories they choose to visit, or indeed not visit, a practitioner of orthodox or alternative medicine; follow or not follow their advice; self-medicate; as well as offer advice to friends and relations. These ideas change over time as a consequence of receiving further information as well as feedback on their behaviors. Preferences for a particular type of cure/therapy are a function of many things. These include beliefs about the cause of the problem as well as the perceived efficacy of the cure. There is also evidence it is also crucially associated with the perceived effort required in, and possible psychological pain associated with treatment. The term psychological pain refers here to the distress associated with the treatment process (Furnham, Chapman, Wilson, & Persuad, 2013).

There are a number of very similar psychological terms like implicit beliefs/theory, health beliefs, lay theories, and mindset that are very closely related. Essentially they refer to an implicit belief system of connected ideas about their own and others’ mental and physical health. In this chapter the term lay theory will be used. By *lay* is meant the opposite of expert or scientific. It is the ordinary, everyday, not specifically ‘educated’ beliefs held by most people. By *theory* is meant the interconnectedness of ideas about illness: their causes, their manifestations, and their cures.

The literature in this area is scattered and goes back a long way. Indeed the ‘health beliefs model’ approach has been popular in health education for years

A. Furnham (✉)

Department of Clinical, Educational and Health Psychology,
University College London, 26 Bedford Way, London WC1 0AP, UK
e-mail: a.furnham@ucl.ac.uk

(Helman, 2007). It is a sophisticated model, which attempts to describe how a person's understanding of such things as the severity of an illness and their personal susceptibility to it might influence their subsequent actions. The idea is that help and intervention are best achieved by eliciting and changing a person's belief about their health.

Two general models of lay beliefs about illness have been proposed. Lay theories can belong to the 'medical' model (Rabkin, 1974), which suggests that mental disorders are like any other illness, with symptoms caused by an underlying biological pathology and a treatment which addresses this. The second model is the 'psychosocial' model, which indicates that causes of mental disorder are psychological and environmental. This has positive implications for treatment, as it advocates social and community support rather than hospitalization. However, it has been found that people with beliefs which correspond to this model are less trusting of former psychiatric patients than ex-medical patients (Furnham, 1988). These models can therefore be used to classify lay theories and have wider implications for attitudes toward those with mental illnesses, causal beliefs, and treatment preferences.

The work on lay theories looks at similar issues, however, it is more concerned with people's understanding of cause and cure than trying to describe how these theories lead to a particular set of behaviors (Furnham et al., 2013). There are also studies on actual knowledge of mental illnesses (Furnham, Gee, & Weis, 2016a). A new focus in a related area is the dramatic increase in the research on mental health literacy (MHL). Further, this literature is complemented by the research on mindsets, which focuses on people's beliefs about change and growth (Furnham, 2015). It is more interested in cognitive skills and development than health but distinguishes between the essentially fatalistic entity believers and the more optimistic incremental believers. The idea is that people 'adopt' a mindset, which has important behavioral consequences.

Doctors have known the importance of 'health beliefs' for a long time and there is an extensive literature on the topic. Helman (2007) noted clinicians "should try to discover how patients and those around them view the origin, significance and prognosis of the condition, and also how it affects other aspects of their lives—such as their income or social relationships. The patient's emotional reactions to ill health (such as guilt, fear, shame, anger, and uncertainty) are all as relevant to the clinical encounter as physiological data, and sometimes more so" (p. 153).

A central question for both researchers and practitioners is how individuals, groups, and societies explain the causes of health and illness, and then what they do about it. People try to make sense of their signs and symptoms of ill health. According to Helman (1990), they typically ask the following questions. These are important questions because they may be seen as the start of the development of a lay theory of all illness.

First, "What has happened?" This includes organizing the symptoms and signs into a recognizable pattern, and giving it a name or identity. People begin to try to make sense of symptoms, seeking advice while they do so. Next, "Why has it happened?" This explains the aetiology of the condition. Third, "Why has it

happened to me?” This tries to relate the illness to aspects of the patient, such as behavior, diet, body build, personality, or heredity. There may be all sorts of social and moral issues in answering this question and justifying various behaviors. This can have a powerful impact on the development of a personal lay theory.

Next, “Why now?” This concerns the timing of the illness and its mode of onset, sudden, or slow. After that has been answered people often ask, “What would happen to me if nothing were done about it?” This considers its likely course, outcome, prognosis, and dangers. It is probable that at this stage the theory, insofar as it is developed, impacts on health-related behavior. The next question is, “What are its likely effects on other people (family, friends, employers, and workmates) if nothing were done about it?” This includes loss of income or of employment, and strain on family relationships. Finally, the last question is, “What should I do about it—or to whom should I turn for further help?” Strategies for treating the condition include self-medication, consultation with friends or family, or going to see a doctor.

While not everyone may be expected to ask all of these questions, they provide a useful list to understand how lay theories develop after some personal illness occurs. It should, however, be pointed out that not all lay theories are about personal illnesses that individuals have experienced: we develop lay theories about all sorts of common mental and physical illness from alcoholism to schizophrenia.

At the heart of much of this work is the very concept of medical and physical health. An example of this research may be illustrative. Over 20 years ago, Blaxter (1990) interviewed over 9000 people from the general public on their understanding about health. About 15% could not think of anyone who was ‘very healthy’ and about 10% could not describe what it was like for them to ‘feel healthy’. This inability to describe what it is like to feel healthy was particularly evident in young males, who believed health to be a norm, a background condition so taken for granted that they could not put it into words.

The categories of health identified from the survey findings were: first, *health as not being ill*, meaning a lack of symptoms, no visits to the doctor; “therefore I am healthy”. Second, *health as reserve*: coming from strong family; recovered quickly from operation. Third, *health as behavior*: usually applied to others rather than self; e.g., they are healthy because they look after themselves, exercise, etc. Fourth, *health as physical fitness and vitality*: used more often by younger respondents the male health concept is more commonly tied to ‘feeling fit’, whereas females had a concept of ‘feeling full of energy’ and rooted health more in the social world in terms of being lively and having good relationships with others. Fifth, *health as function*: the idea of health as the ability to perform ones duties, that is, being able to do what you want when you want without being handicapped in any way by ill health or physical limitation.

People thus have complicated ideas about physical and mental health. These theories develop and change over time, but most importantly inform behavior (Furnham, 1988). Indeed, this topic has attracted a great deal of attention by psychologists, psychiatrists, and others.

Attitudes to Mental Illness, MHL and Lay Theories

Over the years there are three slightly different but overlapping research traditions with regard to this topic. The first concerns studies of attitudes toward people with mental disorders (Nunnally, 1961), that is, beliefs about what people with mental illness are like and how they should be treated. These studies are about specific mental disorders, such as schizophrenia and depression, or more generally about mental illnesses. These are usually large survey-based studies typical of market research or attitudinal surveys. They can offer an explanation for negative and stigmatizing attitudes toward mental illness (and for why so few of those afflicted seek help; Andrews, Hall, Teesson & Henderson, 1999). A recent review looks at public attitudes to a range of disorders from alcoholism and anorexia nervosa to paraphilias, and phobias (Furnham & Telford, 2012). What most of these studies show is that people tend to be ignorant about and often fearful of mental illnesses. It is difficult to summarize this vast area of continuing research stretching back well over 80 years. Few studies have been replicated, however, there are research groups that have done important, systematic, and thorough work on large samples (Angermeyer, Daumer, & Matschinger, 1993). Most of this work is on attitudes to mental illness, while this chapter is concerned more with lay theories.

The second approach is MHL research introduced by Jorm and colleagues (Jorm, 2012). Most of the studies are based on large representative populations, who undergo a structured survey, often by telephone. The group has been particularly interested in depression and schizophrenia, and the perceived pathways of treatment. Other researchers like Furnham have been particularly interested in mental health literature with respect to personality disorders (Furnham, Abajian & McClelland, 2011; Furnham & Wincelhaus, 2012), and more recently conduct disorders (Furnham & Carter Leno, 2012), with results indicating low awareness of the cause, manifestation and cure of the personality disorders.

Studies in this area are also concerned with perceptions of interventions to improve MHL (Smith & Shochet, 2011). Indeed the motivation of most researchers in this area is first to establish the MHL of particular groups and then increase it (Jorm, 2012). Clearly, it is important to increase awareness of MHL as otherwise it may hinder public acceptance of evidence-based mental health care. There is a separate literature on how, when, where, and why people seek help for their, their friend's and their relative's mental health problems. It looks at their knowledge and beliefs about professional as well as self-help and seems always strongly linked to the MHL and the culture from which they come (Jorm, 2012; Sheikh & Furnham, 2012). That requires a separate review.

The third area is the lay theories approach. Recent studies into lay theories have focused specifically on beliefs about the causes and treatments of mental disorders and the relationship between them (Furnham & Buck, 2003), in order to find possible links between negative attitudes and erroneous beliefs. They have revealed that lay theories are not arbitrary or incoherent; they can be classified into categories such as 'psychological' or 'social' in the same way as academic theories (Furnham

& Rees, 1988; Furnham & Thomson, 1996). This suggests that lay people have a basic, possibly implicit, understanding of the different levels of explanation for mental disorders. Studies have shown that the structure of the categories of lay and academic theories overlap to a certain extent, for example, 'biological' and 'psychological' models. Nevertheless some lay conceptions may differ, as regards for example external influences, which include beliefs about the roles of luck and religion in the aetiology of mental illness (Furnham & Buck, 2003).

Lay Theories

Lay theories in psychology and psychiatry are generally thought of as explanations and descriptions which lay people give for various disorders. They have been found to be ambiguous, inconsistent, and incoherent, when compared to academic theories (Furnham, 1988). Research exploring lay beliefs of psychological problems highlights the importance of studying perceived causes and treatments, as this allows an insight into the cognitive strategies people employ when experiencing a problem, or advising both friends and family. Pathways to seeking professional help have been shown to be strongly related to lay theories (Furnham 1988). People give others advice as well as seek it themselves depending on what they think has caused their problem. Thus, they seek medical cures if they believe the cause of their problem is physical, whereas they would favor a 'talking cure' if they believe the problem has affective or cognitive origins. As a result, lay theories could assist in enhancing the effectiveness of some psychological therapies (cognitive behavioral therapy, interventions, and psychotherapy). That is, as the health belief model suggests, one of the first and most important steps in therapy is to elicit and then attempt to change beliefs about illness. These may affect how an individual self-medicates or advises others to do so. They also relate to how optimistic or pessimistic they are about change itself and the possibility of an effective cure.

Initially, research into lay theories focused on investigating lay beliefs about mental illness in general; yet, it seems unlikely that beliefs about diverse illnesses such as schizophrenia and depression are similar (Furnham & Bower, 1992). This has led researchers to investigate beliefs and attitudes concerning different disorders separately. Past research into lay theories has focused on well-known psychological disorders reviewed comprehensively by Furnham and Telford (2012).

When it comes to lay theories, people take very different positions, stressing some factors more than others for both cause and cure. Various studies have identified four types of theories based on the perceived cause of a problem:

The first are biological/genetic/physiological theories for the cause and cure of problems (Furnham, 1988). Some people think there is a chemical imbalance that can be rectified or a genetic abnormality that cannot. Others think that the cause is due to a medical/physical accident (Furnham, 1988). Many see this as the 'modern scientific' approach to understanding mental illness, hoping that neuroscience, behavior genetics, and related new methods will offer new insights into old

problems. Some embrace these types of explanations because they absolve people from taking responsibility for their illnesses, while others like the idea of a physical cause, because it suggests to them that there may be a medical cure. These explanations tend to be favored more by educated, young people from the developed world (Helman, 2007).

Second, some people adopt a psychological/psychoanalytic approach, trying to understand illness in terms of early social experiences or personality factors (Furnham, 1997). The range of theories in this group is great and may be informed by all sorts of psychological ideas. What remains surprising for many is the ‘long reach’ of Freudian ideas that have never been empirically supported: that is, that many Freudian ideas are accepted with little or no evidence (Furnham, 2010). Thus, some people stress early child trauma, poor mothering, or even Freudian personality types (e.g., oral, anal) to explain such issues as alcohol or nicotine addiction. However, even if they do not see the cause of a problem as psychological, many people see ‘the talking cures’ as effective ways to manage or cure the problem.

Third, some adopt a sociological or environmental approach seeing the problem/illness as essentially external to the individual and in the environment. For instance, they express Modern Health Worries, where people seem particularly concerned by things like environmental pollution that they think cause illness (Furnham, 2015). It is not uncommon for people to blame stressful conditions at work for illness, though some will see whole subcultures as pathogenic: that is, they believe that the pattern of behaviors pre- and pro-scribed in a culture can themselves lead to, and maintain various mental illnesses. To some extent, this is the public health approach to illness.

Fourth, still others suggest it is luck, chance, or the will of God. This is essentially a fatalistic approach where people would remark things like “the bullet had your name on it”; “such is life” or “Deus volente”. In some developing countries, witchcraft and evil spirits are quickly evoked as an explanation for illness, and to some extent it is suggested that once specific rituals are performed the ‘spirits’ may be removed.

Lay people usually seem to place more emphasis on psychological, social, and familial cause factors (Angermeyer & Matschinger, 1996a, b; Furnham & Thompson, 1996). Often, they seem to neglect or underestimate the possibility of genetic and biological factors. What is noticeable over the past few decades is how academic theories of the cause and cure of many illnesses have become much more focused on biology, genetics, and physiology. This is partly because of great strides in behavior genetics and modern medicine.

The literature on lay theories may be best illustrated by showing a questionnaire used in lay theory studies (Furnham, Ritchie & Lay, 2016b), which is depicted in Table 1. This list has been used in a number of studies and was derived from many interviews with lay people. It shows the range of ‘causes’ they have suggested for the origins of depression. In these studies, people rate each cause and cure, and these ratings are later factor analyzed to look for underlying patterns.

Using the above questionnaire, Furnham et al. (2016a, b) found the causal questions loaded onto seven interpretable factors labeled God/fate, Environmental,

Table 1 A questionnaire about the causes and cures of depression

Depression is caused by ...	Strongly disagree	2	3	4	5	6	Strongly agree
Depression is a mental disorder characterized by a pervasive low mood loss of interest in usual activities and diminished ability to experience pleasure. Below you will find a list of potential causes of depression. For each item, we would like you to rate how strongly you believe it is a <i>cause of depression</i>							
Genetic factors	1	2	3	4	5	6	7
Being controlled by Satan	1	2	3	4	5	6	7
Complications before or during birth	1	2	3	4	5	6	7
A chemical imbalance in the brain	1	2	3	4	5	6	7
Stress at work	1	2	3	4	5	6	7
Academic pressure or failure	1	2	3	4	5	6	7
Evil done in a previous life	1	2	3	4	5	6	7
Being raised by parents or guardians who have depression	1	2	3	4	5	6	7
A stressful family environment	1	2	3	4	5	6	7
Thinking about things too much	1	2	3	4	5	6	7
Not following religious commandments	1	2	3	4	5	6	7
Day-dreaming too much	1	2	3	4	5	6	7
Enlargement of certain areas of the brain	1	2	3	4	5	6	7
A brain neurotransmitter dysfunction	1	2	3	4	5	6	7
Destiny	1	2	3	4	5	6	7
Cold and uncaring parents	1	2	3	4	5	6	7
Possession by ghosts, genies, or evil spirits	1	2	3	4	5	6	7
Having blood relatives who have depression	1	2	3	4	5	6	7
Taking illegal drugs	1	2	3	4	5	6	7
Searching too much for inner peace	1	2	3	4	5	6	7

(continued)

Table 1 (continued)

Depression is caused by ...	Strongly disagree			Strongly agree		
	1	2	3	4	5	6
A lack of sleep	1	2	3	4	5	7
The pressures of modern society	1	2	3	4	5	7
Punishment from God	1	2	3	4	5	7
Having an overprotective mother	1	2	3	4	5	7
Immoral behavior	1	2	3	4	5	7
Not drinking enough water	1	2	3	4	5	7
Germes or a virus that affects the brain	1	2	3	4	5	7
Loneliness or a lack of friends	1	2	3	4	5	7
Being controlled by a witch doctor	1	2	3	4	5	7
Life trauma (e.g., separation, or loss of a relative/close friend)	1	2	3	4	5	7
Childhood trauma (e.g., physical or sexual abuse)	1	2	3	4	5	7
A monotonous and mundane life	1	2	3	4	5	7
Old age	1	2	3	4	5	7
Being the victim of black magic	1	2	3	4	5	7
An unhealthy diet	1	2	3	4	5	7
A lack of exercise	1	2	3	4	5	7
Smoking too much	1	2	3	4	5	7
A lack of freedom in society	1	2	3	4	5	7
Individuals wanting to be different	1	2	3	4	5	7
Repressed feelings and emotions in the subconscious	1	2	3	4	5	7
Wind	1	2	3	4	5	7
Body temperature	1	2	3	4	5	7
A test from God	1	2	3	4	5	7

(continued)

Table 1 (continued)

Depression is caused by ...	Strongly disagree					Strongly agree				
	1	2	3	4	5	6	7	8	9	10
Financial problems	1	2	3	4	5	6	7			
Lack of will power	1	2	3	4	5	6	7			
A side-effect of some other illness	1	2	3	4	5	6	7			
Unknown causes	1	2	3	4	5	6	7			
Depression is caused by...	Strongly disagree					Strongly agree				
Below you will find a list of potential <i>treatments for depression</i> . For each item, we would like you to rate how strongly you believe it is a good treatment for depression										
Taking prescribed medication or drugs	1	2	3	4	5	6	7			
Acupuncture	1	2	3	4	5	6	7			
Family counseling	1	2	3	4	5	6	7			
Being more religious	1	2	3	4	5	6	7			
Eating well	1	2	3	4	5	6	7			
Seeing a psychiatrist	1	2	3	4	5	6	7			
Becoming more physically active	1	2	3	4	5	6	7			
Finding new friends	1	2	3	4	5	6	7			
Cognitive behavioral therapy (CBT)	1	2	3	4	5	6	7			
Stress management	1	2	3	4	5	6	7			
Quitting illegal drugs	1	2	3	4	5	6	7			
Seeing a faith healer	1	2	3	4	5	6	7			
Being kind to others	1	2	3	4	5	6	7			
Changing one's diet	1	2	3	4	5	6	7			
Being admitted to a mental hospital	1	2	3	4	5	6	7			

(continued)

Table 1 (continued)

Depression is caused by...	Strongly disagree					Strongly agree				
	1	2	3	4	5	6	7	8	9	10
Buying books on depression	1	2	3	4	5	6	7			
Drinking coconut water	1	2	3	4	5	6	7			
Thinking positively	1	2	3	4	5	6	7			
Seeing a counselor	1	2	3	4	5	6	7			
Dealing with symptoms on one's own	1	2	3	4	5	6	7			
Going for a physical check-up	1	2	3	4	5	6	7			
Getting help from close friends	1	2	3	4	5	6	7			
Talk therapy	1	2	3	4	5	6	7			
Resting more	1	2	3	4	5	6	7			
Seeing a social worker	1	2	3	4	5	6	7			
Taking herbal medicine	1	2	3	4	5	6	7			
Being strong emotionally	1	2	3	4	5	6	7			
Quitting smoking	1	2	3	4	5	6	7			
Meditation or yoga	1	2	3	4	5	6	7			
Seeing a GP or doctor	1	2	3	4	5	6	7			
Prayer	1	2	3	4	5	6	7			
Rigorous exercise	1	2	3	4	5	6	7			
Hypnosis	1	2	3	4	5	6	7			
Seeing a psychologist	1	2	3	4	5	6	7			
Taking up a hobby	1	2	3	4	5	6	7			
Drinking more water	1	2	3	4	5	6	7			
Electroconvulsive therapy (ECT)	1	2	3	4	5	6	7			
Exorcism	1	2	3	4	5	6	7			

(continued)

Table 1 (continued)

Depression is caused by...	Strongly disagree			Strongly agree			
	1	2	3	4	5	6	7
Going on holiday	1	2	3	4	5	6	7
Following religious commandants	1	2	3	4	5	6	7
Homeopathy	1	2	3	4	5	6	7
Seeing a psychoanalyst	1	2	3	4	5	6	7
Using telephone counseling service (e.g. Befrienders)	1	2	3	4	5	6	7
Socializing more	1	2	3	4	5	6	7
Getting help from one's close family	1	2	3	4	5	6	7
Seeing a priest or religious teacher	1	2	3	4	5	6	7
Taking some time off work	1	2	3	4	5	6	7
Not doing anything	1	2	3	4	5	6	7

Health, Self-Obsession, Brain, Genetics, and Parents. Similarly, the cure items loaded onto seven factors: Talking Cure, Lifestyle Change, Social Support, Faith, Alternative Medicine, Self-help, and Medical. The cause and cure factors were logically correlated.

Studies in this area tend either to ask people to rate statements about the cause, consequence and cure of specific problems/disorders and/or get them to evaluate vignettes which represent stereotypic problems. These vignettes are usually based on DSM criteria but a variety of different vignettes supposedly representing the same problem do exist.

For instance, in their study Furnham et al. (2016a, b) also used vignette methodology more commonly used in the MHL research.

Case 1:

Siti is 30 years old. She has been feeling really down for the last few weeks. She does not enjoy things the way she normally would. In fact, nothing gives her pleasure. Even when good things happen, they do not seem to make Siti happy. She has to force herself to get through the day, and even the smallest things seem hard to do. She finds it hard to concentrate on anything and has no energy at all. Even though Siti feels tired at night, she still cannot sleep, and wakes up too early in the morning. Siti feels worthless and feels like giving up. Her family has noticed that she has not been herself for about the last month. She does not feel like talking and is not taking part in things like she used to.

In your opinion, what is wrong, if anything, with Siti?

How do you think Siti could best be helped?

Case 2:

Johan is 45 years old. In the recent month, Johan has been feeling unusually sad and miserable. He does not enjoy being with his friends and family as before. Even though he feels tired every day, he found it difficult to sleep at night and struggles to get out of bed in the mornings. He does not feel like eating and has lost a lot of weight. Johan cannot concentrate on his daily tasks, and finds it very difficult to function in the home and at work. Johan cannot keep his mind on his work and puts off making important decisions. This is causing problems in his job and his boss is concerned about his lowered productivity. Johan thinks he is a burden to his family and believes that they would be better off without him. Johan feels so strongly that he is unable to cope with life and unable to be happy anymore, he has been thinking of ways to end his life.

In your opinion, what is wrong, if anything, with Johan?

How do you think Johan could best be helped?

Although there were differences in the responses to the two vignettes and a variety of diagnostic terms used (e.g., feeling down, low confidence), it seemed clear that participants could identify depression. Interestingly, around half suggested some talking cure (counseling/therapy) and half medical intervention (visit doctor, take antidepressants).

Jorm et al. (1997) found that schizophrenia was more likely to be attributed to genetic factors than depression, and lay theories of autism were more likely to be biological than theories of obsessive-compulsive disorder, which were more likely to be psychodynamic in orientation (Furnham & Buck, 2003). It may be that more 'odd', unpredictable, or bizarre the behavior, the more it is likely to be attributed to biological factors. However, the finding that lay theories are generally psychosocial rather than biological has been frequently replicated (Furnham, 1988). This seems the very opposite to the theories held by those in the medical and psychiatric profession (Harland et al., 2009). This is no doubt a function of their training and the search for biological explanations for mental illnesses.

Some individuals are apparently happy to attempt to explain both the cause and best treatment for nearly all mental afflictions by one or another model, while others are quite happy to explain one problem (i.e., addiction or depression or psychosis) by one model or theory, and yet another by another model. So they might offer a biological or physiological explanation for one and a sociological explanation for another. Inevitably, people also 'mix' their models, believing that most causes are multifaceted: that is, that there is more than one type of causal factor acting at the same time. More sophisticated individuals have an interactive approach, believing that psychological and environmental factors can and do interact (Furnham, 1988, 2010).

One recent paper that examined psychiatrists' conceptions of mental illness suggested that eight different models could be described: biological, cognitive, behavioral, psychodynamic, social realists, social constructionist, nihilist, and spiritual (Harland et al., 2009). The biological model stresses genetic and physiological factors; the cognitive model thinking patterns and styles; the behavioral model the idea that illness is cause by maladaptive behaviors; psychodynamic by unconscious processes; social realists stress social factors like poverty, poor housing, and unemployment; social constructionists the idea that disorders are culturally relative and 'made-up' by various groups; nihilists that mental health professionals are charlatans and that no real scientific knowledge exists; while finally some believe there are religious and spiritual explanations for mental illnesses.

Of course psychiatrists are not lay people, being highly educated in many aspects of mental health and illness. However, it is also clear that they differ among themselves often dramatically with regard to all aspects of mental illness: aetiology, classification, definition, intervention, etc.

Treatment and Cure

The general public (as potential clients) is increasingly faced with a bewildering array of psychotherapy interventions available, although some are clearly similar in theory and practice. These include seeing a therapist, attending training courses or focus groups, observation and/or taking medication, or getting hypnosis. Deciding

whether or not to seek help is associated with a range of factors including the availability of services, financial costs, and individual socio-demographic and psychological variables.

Public perceptions of psychotherapists and the process of psychotherapy have been speculated to have important implications in terms of the number and type of individuals who choose to seek psychological treatment (Furnham & Wardley, 1990; Halgin, Weaver & Donaldson, 1985; Wong, 1994). In addition to these influences on potential clients and on their actual experience of treatment, popular perceptions of psychotherapy are likely to have significant implications for public policy and mental health reform (e.g., Knapp & Kamin, 1993; Pallak & Kilburg, 1986). That is, certain experiences in a particular country or region might lead people to call for government-sponsored treatments. Therefore, obtaining a more thorough understanding of the nature of popular perceptions and their antecedents may be helpful in designing interventions to modify negative attitudes towards seeking help (Fischer & Turner, 1970).

The groups least likely to utilize mental health services are men, older people, and people from ethnic minorities, who are all more likely to display avoidance behavior, resistance to treatment and denial of mental illness (Leong & Zachar, 1999). Aside from these influential factors, two major criteria that lay people factor into their choice or recommendation of a therapy presumably are the perceived efficacy of the treatment and the associated side effects for specific psychological issues. For example, counseling is frequently considered most helpful (McKeen & Corrick, 1991) and expectations of counseling involve talking to an experienced expert who can be trusted (Tinsley & Harris, 1976). Prospective patients of many of the talking therapies, particularly psychoanalytic therapies, often seem ignorant of the psychic effort that they are required to make, and the possible emotional pain that results from their therapy (Furnham, 2010). It is expectations such as these that facilitate or hinder the effectiveness of therapy (Apfelbaum, 1958) as well as the choice of therapy.

Various studies have addressed lay beliefs about the best cure for, and ways of overcoming psychological problems (Knapp & Karabenick, 1985). Together, they replicated the factor structure (the number and description of the factors) and cure-specific perceptions of the efficacy of different cures (Furnham & Henley, 1988; Henley & Furnham, 1988) and addictions (Furnham & McDermott, 1994), emphasizing the importance of self-control and, to a lesser extent, professional help, depending on the nature of the disorder. In a series of three studies, Furnham and Wardley (1990, 1991, 1992) investigated lay people's theories regarding the efficacy of various psychotherapy interventions and the prognosis of different disorders. They identified an interpretable underlying factor structure, with lay people discriminating quite clearly between the efficacy of 22 different therapies. It was further found that participants felt largely optimistic about the influence of psychotherapy on various psychological problems, and participant age and education were significant predictors of these beliefs. As often found, younger, better educated people are more positive about the effect of psychotherapy.

One factor that was predictably related to lay theories about psychotherapy was participants' direct or indirect knowledge (through reading) of psychological ideas and therapies. The more experience of treatment that people had, the more skeptical they were about the usefulness of various treatments. Moreover, Furnham, Wardley and Lillie (1992) found that, when compared to lay adults, psychotherapists and students were more skeptical and pessimistic about the efficacy of therapy and prognosis for many illnesses. Knowledge about psychological cures led to a greater awareness of the limited benefits of therapy. However, this finding was not replicated by Furnham (2009) in his investigation of lay attitudes toward and understanding of psychotherapy in treating two psychotic (bipolar, schizophrenia) and two neurotic (depression, obsessive-compulsive) disorders. It is possible that people tend to hold different lay theories about the prospects of treatment of psychotic and neurotic disorders. It was confirmed, however, that participants were generally positive about the experience of psychotherapy but seemed naïve about the actual proven efficacy of psychotherapy.

Many studies have looked at the relationship between lay theories and education in the health and social sciences, as well as personal experiences of illnesses (Furnham, 1988; Helman, 2007). It is usually hypothesized that better educated people would be more accurate about the prognosis of particular illnesses as well as the efficacy of different therapeutic interventions. However, this hypothesis has not always been confirmed showing either no relationship between education and lay knowledge and even occasionally a negative correlation (Furnham, 1988) This may be due to experimental methods such as not getting enough or sensitive information on personal on knowledge.

People are often faced with a wide array of therapeutic interventions from drugs to talking therapies. The list differs depending on the nature of the problem to the availability of the therapy. Psychotherapies involving cognitive, affective, and behavioral procedures have been established as empirically supported treatments for anxiety disorders (e.g., Chambless & Ollendick, 2001). Most cognitive and behavioral techniques are derived from theoretically coherent and empirically validated models of anxiety disorders, and provide a consistent relationship between the treatment techniques and symptoms. In addition, from the client's perspective, they have fewer (at least biological and pharmacological) undesirable side effects, and no addictive potential that may occur with other therapies, including medication.

In many anxiety disorders clients, particular situations are avoided because of the aversive affect they produce: this is judged by the patients themselves and professionals as unmanageable and requiring treatment. All therapists accept the idea that affects needs to be faced (through behavioral exposure or talking about past events) and that, in doing so, it becomes much more manageable. Developments in radical behaviorism are particularly strong on this idea (Hayes, Strosahl, & Wilson, 1999; Linehan, 1993). Exposure to affect has also always been central to the psychodynamic and humanistic therapies, with usually beneficial results (Bornstein, 1989).

Generally, most members of the public believe that mental disorders are treatable, and psychiatric treatments are considered generally rather unhelpful whereas counseling is considered most helpful (Furnham et al., 2001). Studies have also shown that people have set ideas about counseling before taking up therapy. Expectations have been found to be important determinants of where people turn to for help and effectiveness of counseling (Furnham, 1988; Furnham et al., 2013). Many people underestimate how difficult, effortful, and painful it is to undergo various therapeutic treatments preferring those that seem quick and easy, and which are often ineffective (Furnham et al., 2013). That is, they seek a 'magic pill' for a complex problem. Indeed, it is similar to the issues in the diet industry that shows people seek a dramatic short-term diet to lose weight but soon resume a lifestyle that caused the problem in the first place. Therapy, if successful, often involves considerable effort and pain as people adjust to a new lifestyle. Yet this is often misunderstood by lay people whose understanding is often based on very inaccurate reports.

People have very different beliefs about what occurs during psychotherapy (Furnham & Telford, 2012). Furnham and Wardley (1990) found respondents tended to believe that clients of psychotherapy did feel better in therapy, and were more confident and hopeful. These results were replicated by Wong (1994) and Heaven and Furnham (1994) using American college students and staff. Knowledge about psychological cures led to a greater awareness of the limited benefits of therapy. This was confirmed when Furnham et al. (1992) compared responses of lay adults, students, and clinical psychologists, and found the latter tended to be more cynical about the efficacy of therapy and prognosis of many disorders.

There is also a belief that 'will power' can effectively facilitate recovery from mental disorders (Knapp & Delprato, 1980), such as agoraphobia and anorexia nervosa (Furnham & Henley, 1988). However, medication is believed to be the most effective treatment for disorders with a higher perceived severity (Furnham & Bower, 1992; Furnham & Rees, 1988), thus showing that lay and academic theories of treatment overlap to an extent.

The question of 'will power' is interesting and not well explored in the literature. It may be thought of as the desire and ability to adhere to treatment recommendations which could be seen as tedious or difficult. Certainly with regard to addictions (alcohol, food, tobacco) lay people say will power is an essential component. Hence the joke: how many psychologists does it take to change a light bulb? Only one; but the light bulb needs to want to be changed.

The Relationship Between Cause and Cure

Studies have also focused on assessing whether there is a logical relationship between lay theories of cause and treatment. It is expected that if the cause is attributed to biological factors, medication should be endorsed as treatment, and if the cause is psychosocial some behavioral or psychological treatment may be

recommended. This has been found in a number of studies, which show a strong relationship between similar cause and treatment theories (Furnham & Buck, 2003), and those which are ‘sensibly’ linked (Furnham & Haraldsen, 1998, pp. 696). These findings are not always replicated, for example, medication was the preferred treatment for schizophrenia, despite participants attributing the cause to psychosocial factors (Furnham & Bower, 1992; Furnham & Rees, 1988). This may well be due to the acuity, chronicity, and severity of the symptoms. On the other hand, there are some who are highly critical of the use of psychopharmacology even in the most extreme psychiatric disorders (e.g., Breggin & Cohen, 2007), in large part because of the potentially adverse effects of such medication.

In sum, studies in this area have covered a very wide range of specific illnesses including alcoholism, anxiety, anorexia, autism, depression, gender identity disorder, heroin addiction, neurosis, paraphilia, phobia, schizophrenia, and suicidal thoughts (Furnham & Telford, 2012). Many have looked at what lay people think causes these particular problems and how they are behaviorally manifested as well as the most effective treatment modality (Furnham & Telford, 2012). They show areas of ignorance about mental illnesses, particularly developments in isolating and describing bio-physical and genetic processes relating to those illnesses. Essentially, they demonstrate that lay people take more of a psychological than a biological/physiological approach to the cause of a wide range of illnesses, and are (as may be expected) not up-to-date with scientific advances in many areas.

Conclusion

Over time people develop ‘theories’ as to the causes and consequence of their mental and physical health. These theories develop and change over time as function of personal experience. They are related to a number of factors including the culture in which a person lives and grows up, their primary and secondary socialization, and their personality and values. People learn from personal experience and are exposed to numerous ideas about the causes and cures of various problems which they try to integrate into a theory.

More importantly, lay theories influence social behavior. Thus people may avoid foods or social situations because they believe them to be toxic. Equally, they may seek out people and medicines because they feel they are beneficial for their mental and physical health. They are encouraged by family and friends, as well as range of experts to change their behavior in order to become more happy and healthy.

The bio-psycho-social approach to (mental and physical) health and illness suggests that health is a function of biological and psychological and social factors. Lay theories are dynamic cognitions like health beliefs which shape a person’s behavior. To some extent it is possible to see cognitive behavior therapy as a way of trying to improve mental health by changing attributions and hence theories of other people and other situations (Furnham et al., 2013). For years, clinical and health

psychologist have been interested in the health beliefs model; equally personality and social psychologist have been most concerned with lay beliefs which they feel equally important in shaping a person's mental health.

References

- Andrews, G., Hall, W., Teesson, M., & Henderson, S. (1999). *The mental health of Australians*. Canberra: Commonwealth Department of Health and Aged Care.
- Angermeyer, M. C., Daumer, R., & Matschinger, H. (1993). Benefits and risks of psychotropic medication in the eyes of the general public: Results of a survey in the Federal Republic of Germany. *Pharmacopsychiatry*, *26*, 114–120.
- Angermeyer, M. C., & Dietrich, S. (2006). Public beliefs about and attitudes towards people with mental illness: A review of population studies. *Acta Psychiatrica Scandinavica*, *113*, 163–179.
- Angermeyer, M., & Matschinger, H. (1996a). The effect of diagnostic labeling on the lay theories regarding schizophrenic disorders. *Social Psychiatry and Psychiatric Epidemiology*, *31*, 316–320.
- Angermeyer, M. C., & Matschinger, H. (1996b). Public attitude towards psychiatric treatment. *Acta Psychiatrica Scandinavica*, *94*, 326–336.
- Apfelbaum, D. (1958). *Dimensions of transference in psychotherapy*. Berkeley: University of California Press.
- Blaxter, M. (1990). *Health and lifestyle*. London: Tavistock/Routledge.
- Bornstein, R. (1989). Exposure and affect: Overview and meta-analysis of research 1969–1987. *Psychological Bulletin*, *106*, 265–289.
- Breggin, P. R., & Cohen, D. (2007). *Your Drug May Be Your Problem: How and Why to Stop Taking Psychiatric Medications, Second Edition*. Cambridge MA: Perseus Books.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685–716.
- Fischer, E. H., & Turner, J. L. (1970). Orientations to seeking help: Development and research utility of an attitude scale. *Journal of Consulting and Clinical Psychology*, *35*, 79–90.
- Fischer, W., Goerg, D., Zbinden, E., & Guimon, J. (1999). Determining factors and the effects of attitudes towards psychotropic medication. In J. Guimon, W. Fischer, & N. Sartorius (Eds.), *The image of madness: The public facing mental illness and psychiatric treatment* (pp. 162–186). Basel: Karger.
- Furnham, A. (1988). *Lay theories: Everyday understanding of problems in the social sciences* (p. 252). Oxford: Pergamon.
- Furnham, A. (1997). Overcoming neurosis: Lay attitudes of cure for five specific neurotic disorders. *Journal of Clinical Psychology*, *53*, 595–604.
- Furnham, A. (2009). Psychiatric and psychotherapeutic literacy: Attitudes to, and knowledge of, psychotherapy. *International Journal of Social Psychiatry*, *55*, 525–537.
- Furnham, A. (2010). Psychiatric and psychotherapeutic literacy: Attitudes to, and knowledge of, psychotherapy. *International Journal of Social Psychiatry*, *55*, 525–537.
- Furnham, A. (2015). What you can and can't change: Lay perspectives on Seligman's Guide. *Psychology*, *5*.
- Furnham, A., Abajian, N., & McClelland, A. (2011). Psychiatric Literacy and the Personality Disorders. *Psychiatry Research*, *189*, 110–114.
- Furnham, A., & Bower, P. (1992). A comparison of academic and lay theories of schizophrenia. *British Journal of Psychiatry*, *161*, 201–210.
- Furnham, A., & Buck, C. (2003). A comparison of lay-beliefs about autism and obsessive-compulsive disorder. *International Journal of Social Psychiatry*, *49*, 287–307.

- Furnham, A., & Carter-Leno, V. (2012). Psychiatric literacy and the conduct disorders. *Research in Developmental Disabilities, 33*, 24–31.
- Furnham, A., Gee, M., & Weis, L. (2016a). Knowledge of mental illness. *Psychiatry Research, 244*, 363–369.
- Furnham, A., & Haraldsen, E. (1998). Lay theories of aetiology and “cure” for four types of paraphilia: Fetishism, paedophilia, sexual sadism and voyeurism. *Journal of Clinical Psychology, 54*, 689–700.
- Furnham, A., & Henley, S. (1988). Lay beliefs about overcoming psychological problems. *Journal of Social and Clinical Psychology, 26*, 423–438.
- Furnham, A., & McDermott, M. (1994). Lay beliefs about the efficacy of self-reliance, seeking help and external control as strategies for overcoming obesity, drug addiction, marital problems, stuttering and insomnia. *Psychology and Health, 9*, 397–406.
- Furnham, A., Pereira, E., & Rawles, R. (2001). Lay theories of psychotherapy: Perceptions of the efficacy of different ‘cures’ for specific disorders. *Psychology, Medicine and Health, 6*, 77–84.
- Furnham, A., & Rees, J. (1988). Lay theories of schizophrenia. *International Journal of Social Psychiatry, 34*, 212–220.
- Furnham, A., Ritchie, W., & Lay, A. (2016b). Beliefs about the causes and cures of depression. *International Journal of Social Psychiatry, 62*, 415–424.
- Furnham, A., & Telford, K. (2012). Lay theories of, and public attitudes towards, mental illness. In Mental Health (Ed.), *L’Abate* (pp. 3–22). Intech: Rijeka.
- Furnham, A., & Thomson, L. (1996). Lay theories of heroin addiction. *Social Science and Medicine, 43*, 29–40.
- Furnham, A., & Wardley, Z. (1990). Lay theories of psychotherapy I: Attitudes toward, and beliefs about, psychotherapy and therapists. *Journal of Clinical Psychology, 46*, 878–890.
- Furnham, A., & Wardley, Z. (1991). Lay theories of psychotherapy II: The efficacy of different therapies and prognosis for different problem. *Human Relations, 44*, 1197–1211.
- Furnham, A., Wardley, Z., & Lillie, F. (1992). Lay theories of psychotherapy III: Comparing the ratings of lay persons and clinical psychologists. *Human Relations, 45*, 839–858.
- Furnham, A., Chapman, A., Wilson, E., & Persuad, R. (2013). Treatment hurts: Lay theories of graded exposure in the treatment of four anxiety disorders. *European Journal of Psychotherapy and Counseling, 15*, 253–273.
- Furnham, A., Kirkby, V., & McClelland, A. (2011). Non-expert’s theories of three major personality disorders. *Personality and Mental Health, 5*, 43–56.
- Furnham, A., & Wincelous, J. (2012). Psychiatric literacy and the personality disorders. *Psychopathology, 45*, 29–41.
- Halgin, R. P., Weaver, D. D., & Donaldson, P. E. (1985). College students’ perceptions of the advantages and disadvantages of obtaining psychotherapy. *Journal of Social Service Research, 8*, 75–86.
- Harland, R., Antonova, E., Owen, G., Broome, M., Landau, S., Deeley, Q., et al. (2009). A study of psychiatrists’ concepts of mental illness. *Psychological Medicine, 39*, 967–976.
- Hayes, S., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy*. New York: Guildford Press.
- Heaven, P., & Furnham, A. (1994). Personality factors and preferences for therapy. *Australian Psychologist, 29*, 207–211.
- Helman, C. A. (2007/1990). *Culture, health and illness*. London: Hodder/Arnold.
- Henley, S., & Furnham, A. (1988). The attribution of cure. *British Journal of Clinical Psychology, 27*, 384–386.
- Jorm, A. (2012). Mental health literacy. *American Psychologist, 67*, 231–243.
- Jorm, A., Korten, A., Jacomb, P., Christensen, H., Rodgers, B., & Pollitt, P. (1997). Public beliefs about causes and risk factors for depression and schizophrenia. *Social Psychiatry and Psychiatric Epidemiology, 32*, 1143–1148.
- Knapp, H. D., & Kamin, J. (1993). *Therapy’s value*. APA Monitor, p. 3.
- Knapp, J., & Delprato, D. (1980). Willpower, behavior therapy, and the public. *Psychological Record, 30*, 477–482.

- Knapp, J., & Karabenick, S. (1985). Overcoming problems. *Cognitive Therapy and Research, 9*, 343–354.
- Leong, F., & Zachar, P. (1999). Gender and opinions about mental illness as predictors of attitudes toward seeking professional psychological help. *British Journal of Guidance and Counselling, 27*, 123–132.
- Linehan, M. (1993). *Cognitive behavioural treatment of borderline personality disorder*. New York: Guilford Press.
- McKeen, P., & Corrick, S. (1991). Public attitudes to depression: A national survey. *Irish Journal of Psychological Medicine, 8*, 116–121.
- Nunnally, J. (1961). *Popular conceptions of mental health*. New York: Holt, Rinehart & Winston.
- Pallak, M. S., & Kilburg, R. R. (1986). Psychology, public affairs, and public policy: A strategy and review. *American Psychologist, 41*, 933–940.
- Priest, R. G., Vize, C., Roberts, A., Roberts, M., & Tylee, A. (1996). Lay people's attitudes to treatment of depression: Results of opinion poll for defeat depression campaign just before its launch. *British Medical Journal, 313*, 858–859.
- Rabkin, J. G. (1974). Public attitudes toward mental illness: A review of the literature. *Psychological Bulletin, 10*, 9–33.
- Sheikh, S., & Furnham, A. (2012). The relationship between somatic expression, psychological distress and GP consultations in two cultural groups. *Counselling Psychology Quarterly, 25*, 389–402.
- Smith, C. L., & Shochet, I. M. (2011). The impact of mental health literacy on help-seeking intentions: Results of a pilot study with first year psychology students. *International Journal of Mental Health Promotion, 13*, 14–20.
- Tinsley, H., & Harris, S. (1976). Client expectations for counseling. *Journal of Counseling Psychology, 23*, 173–177.
- Wong, J. L. (1994). Lay theories of psychotherapy and perceptions of therapists: A replication and extension of Furnham and Wardley. *Journal of Clinical Psychology, 50*, 624–632.

Index

A

- Aetiology, 356, 359, 367
 - etiological gap, 290, 291
 - genetic theories, 33
 - of obesity, 326, 327, 330, 331
- Agency detection
 - etiological gaps to, 290–291
 - misperception to, 290
 - motion to, 290
- Algorithmic complexity, 266, 283, 288, 299
- Anthropological challenge, and developmental psychology, 297–299
- Anthropology of religion, 273–276, 306
- Anti-free will group, 234
- Artificial Intelligence (AI), 108
- Attraction-selection-attrition (ASA) framework, 118

B

- Belief
 - in afterlife, 207–208
 - in free will. *See* Free will
 - in God. *See* Big Gods beliefs
 - in mind–body dualism, 196–198
 - on mind/brain relationship and self
 - central, 218–219
 - cohesiveness of, 217
 - concept, 217
 - CORE and PERIPHERAL attributes, 218
 - near vs. distant future, 218
 - students' assessment, 217–218
 - temporal self-appraisal theory, 218
 - terminology, 182
- Belief in a just world
 - active strategies, 162
 - benefits, 160
 - endorsement, 160–161
 - function of, 160

- management, 166
- people's reactions toward unjust events, 159–161
- threat, 162, 163, 165–171
- Big-C creativity, 99
- Big Gods beliefs
 - by-products
 - closer relationship with God, 257
 - evolutionary, 249
 - human morality, 255–257
 - negative, 249
 - prosocial punishment, 254–255
 - protecting humans, 256–257
 - self-regulation, 252–254
 - watchful Gods, 250–251
 - evidence for, 245–248
 - religious beliefs, 246–247
- Bio-psycho-social approach, 371
- Body shame, 329, 333
- Brainstorming session, 111

C

- Cancer
 - and bellicose metaphors, 344–347
 - enemy metaphors for, 349
- Cartesian Theater, 196, 198
- Cognitive science of religion, 266, 288, 289, 309
- Cognitive stimulation, 111, 112
- Combinational creativity, 108
- Common sense belief, 219, 220
- Common sense justice, 158
- Communications, 10
- Compatibilism, 214
- Compensatory rationalizations, 170, 171
- Complexity drop model of supernatural (CDMS)
 - case study, 285–286

- Complexity drop model of supernatural (CDMS) (*cont.*)
 epidemiological vs. doxastic models, 289–290
 vs. HADD model, 290–291
 natural causal schema, 286
 vs. probabilistic models, 291–293
 supernatural causal schemas, 286
 types, 287–288
- Consensual Assessment Technique (CAT), 116
- Core knowledge, 279, 280, 296
- Counterintuitive-magic (CT-magic)
 children's
 conception, 267–268
 judgment of reality vs. magical status, 269
 magical power to supernatural agents, 270–271
 magical power to themselves, 269–270
 preferential learning from magical person, 271–272
 in Christian cultures, 306
 and probabilistic-magic
 across development, 299–304
 developmental psychology, 297–299
 secularization and explanatory coexistence, 293–296
 in Shipibo culture, 306–307
 types, 281–282
 varieties of, 279–282
- Creative cognition approach, 100, 104
- Creative person
 age
 children, 103
 older adults, 103–104
 creative genius theory
 fundamental differences, 99–100
 rare creativity, 98–99
 stable traits, 100
 unlearnable creative, 100–101
 psychopathology
 approach vs. avoidance motivation, 102
 mad genius belief, 101
- Creative process
 convergent thinking, 105–106
 creative thinking skills, 104–105
 divergent thought, 105
 flexibility pathway, 106
 idea generating, 109
 incubation phase, 107
 'ordinary' cognitive, 104
 persistence pathway, 106
 stages, 107
 unexpected, 106–107
 uniquely human, 107–108
- Creative product
 CAT technique, 116
 levels of creativity, 115–116
- Creativity, lay theories of
 and age
 children, 103
 older adults, 103–104
 characteristics, 97
 consequences of
 lack of developmental opportunities, 117–118
 self-selection, 118
 stereotype threat, 116–117
- creative genius theory
 fundamental differences, 99–100
 rare creativity, 98–99
 stable traits, 100
 unlearnable creative, 100–101
- definition, 97
- 'Four P' model, 97
- place
 brainstorming session, 111
 cognitive stimulation, 111, 112
 freedom vs. constraints, 112–114
 grouping creativity, 111–112
 three B's (Bath, Bus, Bed) of creative ideas, 110
 process. *See* Creative process
 product. *See* Creative product
 and psychopathology
 approach vs. avoidance motivation, 102
 mad genius belief, 101
 role, 95
 significance of, 95
 stereotypes, 96
 types of, 99
- Cultural evolution, Big Gods beliefs
 by-products
 closer relationship with God, 257
 evolutionary, 249
 human morality, 255–257
 negative, 249
 prosocial punishment, 254–255
 protecting humans, 256–257
 self-regulation, 252–254
 watchful Gods, 250–251
 evidence for, 245–248
- Cure
 and cause
 questionnaire, 360–365
 relationship between, 370–371
 theories for, 359
 psychological, 369, 370

- talking, 359, 360, 366
and treatment, 367–370
- D**
- Daydreaming, 71
- Decoupled thought, 71
- Depression, 102, 358, 359, 360, 366, 367, 369, 371
questionnaire on causes and cures of, 361–365
- Deservingness, 171–172
- Developmental psychology, 297–299
- Discrimination, 139, 148
definition, 128
fixed and growth mindsets, 137
negative behavior, 131
real-world, 137
- Double-edged sword effects, 328, 329
- Dualism
mind–body. *See* Mind–body dualism
mind/brain relation, 219–220
- Dyslexia, 223
- E**
- Ego depletion experience, 50–53
glucose in, 56
self-control exertion, 52
- Ego depletion effect, 48
in Indians, 63
motivation and attention, 58
- Electroencephalography (EEG), 235
- Electronic brainstorming systems (EBS), 112
- Entity theory, 6, 9, 24, 100
- Equality, 171, 172
- Equity, 128, 149, 158, 171, 172
- Essentialism
genetic, 12
neuro-, 219, 224
psychological, 4–7
- Explorational creativity, 108
- F**
- Fixed mindsets, 128, 130, 136, 138, 141, 144, 146–148
vs. growth mindsets, 128, 132, 141, 143
- Free will, 181, 182, 190, 196, 198, 208
belief in, 209
affecting social behavior, 233–235
and agentive control, 230–232
cognitive and neural mechanisms, 235–238
definition, 230
empirical support, 232–233
individuals', 231
meta-representation, 238–239
moral rules, 232–233
existence of, 229
neuroscience, 214–216
sensory-motor account, 232
- G**
- Generics, 9–10
- Genetic essentialism, 12
- Goal-striving, personal, 54–55
- God
beliefs in. *See* Big Gods beliefs
definition, 244
- Growth mindsets
benefits of, 148
children, 133, 136
definition, 128
vs. fixed mindsets, 128, 132, 141, 143
and intelligence, 140
minorities and women, 143, 144
participant's willingness, 146
role, 137
and stereotyping environment, 141
- H**
- Health
categories, 357
mental
cognitive behavior therapy, 371
problems, 355
services, 368
metaphors
body and machine, 347–349
cancer and bellicose, 344–347
physical, 357, 371
- Health beliefs model approach, 355–356
- Helping, 168, 169, 171
- Historical creativity, 99
- Hypersensitive agency detection device (HADD) model, 290–291
- I**
- Illness, 355
medical model, 356
mental. *See* Mental illness
metaphors of. *See* Metaphors
psychosocial model, 356
- Immanent justice, 170, 171
- Implicit belief system, 355
- Implicit theories
of change and stability
factors affecting, 22–23
self-image goals, 24–29
in children's sports, 29

- Implicit theories (*cont.*)
 entity and incremental theories, 320
 entity mindsets, 28
 epigenetic research, 19
 experimental malleability, 21–22
 fixed vs. malleable nature of human, 328
 future directions, 37–38
 incremental theory, 27–28
 motivated fluidity, 23–24
 people's belief, 17–18
 person perception and stereotyping, 18
 of social mobility, 35–36
 temporal stability, 20–21
 of weight, 324, 325, 329, 333
- Incremental theory, 100, 320–322, 325, 327, 329, 331–333
- Inherence heuristic, 8, 9
- Intentional binding, 231
- Intentional control, 230–232
- Intergroup relations, 147–149
- Intrusive thoughts, 79–85
 and control beliefs, 82–83
 depression, 79
 description, 79
 generalized anxiety disorder, 79
 healthy individuals, 79–80
 OCD, 79, 80
 psychological distress, 81–82
 rebound effects, 80
 thought suppression, 80
 Wegner's ironic monitoring theory, 80–81
- Intuitive beliefs, 183, 192, 200
- J**
- Judgments, 31
 children's, on reality vs. magical status, 269
 creativity, 115
 of crime, 32
 intergroup, 18
 justice, 158, 165–167, 173
 neuroscience influences on, 216, 223
 free will, 214–216
 responsibility, 212–214
 of politician, 30
 thin-slice person, 25
- Justice
 belief in a just world
 active strategies, 162
 benefits, 160
 endorsement, 160–161
 function of, 160
 management, 166
 people's reactions toward unjust events, 159–161
 threat, 162, 163, 165–171
 commonsense, 158
 definition, 157
 deservingness, 171–172
 individual, 157
 motive theory, 159
 perceptions of, 157–158
 procedural, 158
 reaction following unjust events
 compensatory rationalizations, 170, 171
 derogatory and supportive reactions, 169
 immanent justice reasoning, 170, 171
 people's spontaneous reactions, 169
 positive, 168–169
 self vs. other-focus, 169
 ultimate justice reasoning, 170, 171
 research on victim blaming, 162–163
 approach vs. avoidance orientation, 165–166
 experimental vs. rationalistic processing, 166–168
 self vs. others focus, 164–165
 retributive, 159
 rules of, 157
 social research, 158
- Just-world theory, 159, 162, 168, 171
- K**
- “Kind of person” theories, 129
- Kolmogorov's theory, 283
- L**
- Lay belief. *See also* Implicit theories; Justice, 17, 96
- Lay theor(y)ies, 17
 cognitive foundations, 7–9
 cultural settings, 11–13
 descriptions, 72
 developmental experiences, 9–11
 genetic, 12
 intrusive thoughts. *See* Intrusive thoughts
 mental illness. *See* Mental illness
 psychological essentialism, 4–7
 research on, 72
 social arrangements, 13–14
- M**
- Magic
 anthropological
 approaches, 272–273
 case studies, 276–279
 data formalization, 283–292
 debating of boundaries, 273–276

- varieties of counterintuitions, 279–282
- building blocks of, 266
- counterintuitive-magic and probabilistic-magic
 - across development, 299–304
 - anthropological challenge, 297–299
 - developmental psychology, 297–299
 - secularization and explanatory coexistence, 293–296
- impossible events, 265, 266
- improbable ones, 265, 266
- psychological approach
 - children’s conception, 267–268
 - children’s judgment, 269
 - children’s power to other agents, 270–272
 - children’s power to themselves, 269–270
- types and subtypes, 304–308
- Magical thinking, 266, 268, 272, 273, 283, 301, 309
- Mental control, 84
- Mental health literacy (MHL), 356, 358
- Mental illness
 - aetiology of, 359
 - cause and cure, relationship between, 370–371
 - lay theories
 - biological model, 367
 - cause and cure factors, 360, 366
 - cognitive model, 367
 - literature on, 360
 - luck/chance/God’s will, 360
 - modern scientific approach, 359–360
 - nihilist, 367
 - psychodynamic processes, 367
 - psychological/psychoanalytic approach, 360
 - psychology and psychiatry, 359
 - questionnaire, 360–365
 - social constructionists, 367
 - social realists and social factors, 367
 - sociological/environmental approach, 360
 - spiritual explanations, 367
 - vignettes methodology, 366
 - and MHL, 358
 - timing of, 357
 - treatment and cure, 367–370
- Metacognition, 238–239
- Meta-cognitive beliefs, 72
- Metaphoric framing, 345, 346
- Metaphors
 - and cancer, 341
 - bellicose metaphors, 344–347
 - enemy for, 349
 - effects of, 350
 - and lay theories, 342–344
 - machine
 - and body, 347
 - brain model, 349
 - evolution, 348
 - medical discourse, 347
 - people’s beliefs, 347
 - reductionism, 348–349
 - specific functions, 348
 - pervasiveness, 351
 - use in medicine, 349–350
- Mind wandering, 71
 - assessment, 74
 - during day-to-day activities, 74
 - experience-sampling studies, 73
 - self-serving bias, 75
 - timings during, 74
- Mind–body dualism
 - beliefs in, 196–198
 - bodily self-awareness, 192–193
 - causal evidence, 198–199
 - as cognitive default, 185–187
 - dissociation, 192–193
 - explicit beliefs in, 193–196
 - and free-will, 198
 - future research, 200
 - intuitive/implicit beliefs in, 184
 - on other’s minds, 184–185
 - on our own mind, 190–191
 - paranoid thinking, 200
 - philosophical theory of, 182–183
 - self-regulation resources, 199
 - Theory of Mind development, 187–190
- Mind/brain relationship
 - common sense beliefs, 219–220
 - dualism, 219–220
 - historical background, 207
 - lay theories of
 - brain assessment, 208–209
 - education, 222
 - examples, 221
 - implicit approach, 209
 - natural world, 220
 - public policy in mental health, 221
 - questionnaires form, 208
 - scientific conceptualization, 222–223
 - social world, 221
 - and self
 - central, 218–219
 - cohesiveness of, 217
 - concept, 217

- CORE and PERIPHERAL attributes, 218
 near vs. distant future, 218
 students assessment, 217–218
 temporal self-appraisal theory, 218
- Mindfulness meditation training, 84
- Mindsets
 domain specificity, 320–321
 fixed, 320
 goals people set, 322
 growth. *See* Growth mindsets
 implicit theories, 321–322
 weight-based
 cultural variations, 332
 development of, 330–331
 dieting-related products, 324
 implicit theories, 324–325
 incremental theory, 331–332
 low-cost intervention efforts, 324
 management strategies, 323
 and self-regulation, 324–327
 stigma, 327–329
- Mindsets about malleability, 127, 129–130
 characteristics, 128
 domain specific, 128–129
 fixed mindsets, 128
 growth mindsets. *See* Growth mindsets
 implicit theories, 129
 intergroup relations, 131, 147–149
 perceivers' mindset
 categorization, 131–132
 prejudice. *See* Prejudice
 stereotyping. *See* Stereotyping
 targets
 intergroup reconciliation, 144–146
 overt expression of prejudice, 142–144
 social identity threat, 139–142
- Moral responsibility, 209, 213–215, 232–235
- Motivated reasoning, 23, 24, 34, 37
- N**
- Neuro-essentialism, 219, 224
- Neuroimaging, 209, 223, 302
- Neuroscience
 allure of, 209–210, 212
 cognitive, 210
 development of, 209–210
 judgments of
 free will, 214–216
 responsibility, 212–214
 mapping process, 210
 by neuroimaging, 223
 psychological phenomenon, 210–211
 reduced mental capacities, 212–213, 223–224
- Nominal Group Technique (NGT), 112
- O**
- Obesity, 319, 323, 325, 328, 330–333
 weight-based mindsets, 323–324
 and self-regulation, 324–327
 and stigma, 327–329
- Obsessive–compulsive disorder (OCD), 79, 80, 367
- P**
- Physiology, 48, 360
- Post-error adaptation, 238
- Powerful God's by-products
 prosocial punishment, 254–255
 self-regulation, 252–254
- Prejudice, 130
 definition, 127
 discrimination, 137
 intergroup expression, 136, 137
 lay theories of, 137–139
 measures of, 136
 outgroups, 136–137
 targets of, 142–144
- Probabilistic-magic (PR-magic)
 in Christian cultures, 306
 complexity drop
 CDMS. *See* Complexity drop model of supernatural (CDMS)
 as cue to agency, 283–285
 counterintuitive-magic
 across development, 299–304
 developmental psychology, 297–299
 secularization and explanatory
 coexistence, 293–296
 in Shipibo culture, 306–307
 subtypes across cultures, 308
- Probabilistic reasoning, 266
- Pro-C creativity, 99
- Promotion focus, 100
- Prosocial punishment, 254–255
- Psychological-creativity, 99
- Public understanding of science. *See* Lay theor(y)ies
- Punishment, 31–33, 165, 248
- Q**
- Quality of ideas, 109, 111
- Quality of life, 172, 350
- R**
- Relationship

- cause and cure, relationship between, 370–371
- closer relationship with God, 257
- mind/brain relationship and self
 - central, 218–219
 - cohesiveness of, 217
 - concept, 217
 - CORE and PERIPHERAL attributes, 218
 - near vs. distant future, 218
 - students' assessment, 217–218
 - temporal self-appraisal theory, 218
- Religion
 - behavioral component, 244
 - belief in God, 244–245
 - beliefs, 244
 - communities, 244
 - moral code, 244
- Remote Associates Test, 105
- Rest goal, 57–58
- Risk-taking, 256
- S**
- Science attitudes and perception
 - anti-fat attitudes, 333
 - health related, 244, 368
 - towards Lay theory. *See* Lay theor(y)ies
 - towards mental illness. *See* Mental illness
 - towards MHL. *See* Mental health literacy (MHL)
- Scientific communication, 21
- Self-control
 - capacity, 47
 - cross-cultural differences, 63
 - effects
 - everyday self-regulation, 53–54
 - personal goal-striving, 54–55
 - well-being, 55
 - ego depletion effect, 48
 - failure in, 47
 - interventions, 63–64
 - limited resource theory, 50–53
 - measuring for strenuous mental activity, 49–50
 - mechanisms
 - activation of rest goal, 57–58
 - interaction process, 58–59
 - mental resources, availability of, 56–57
 - perceived exhaustion, 55–56
 - self-efficacy, 58
 - negative consequences
 - interpersonal consequences, 60
 - overuse resources, 59–60
 - parenting, 61
 - nonlimited theory, 49
 - social learning theory, 62
 - social-psychological research, 47
 - strength model, 47–48
- Self-determination theory (SDT), 113, 114
- Self-efficacy, 58
- Self-regulation, 53–54, 252–254, 324–327
- Self-selection, 118
- Sense of agency, 230–231
- Social cognitive theory, 323
- Socioeconomic status (SES), 35, 36
- Spontaneous thoughts, 75–76
- Statistical inference, 210
- Stereotyping, 130
 - definition, 127
 - endorsement, 135
 - formation, 133–134
 - women's perceptions of, 141
- Stigma
 - asymmetry model, 328
 - definition, 128
 - weight-based mindsets, 327–329
- Stimulus-unrelated thoughts, 76–78, 83, 85–86
 - controllability of, 76–78
 - intrusive character, 71
 - meaning of, 75–76
- Supernatural explanation, 292, 294–296, 301
- System justification, 8, 14, 34, 36, 171
- T**
- Task-unrelated thoughts, 71
- Theory of Mind, 187–190, 200
- Thought-action fusion, 80
- Thought suppression, 80–82
- Transformational creativity, 108
- Treatment and cure, 367–370
- U**
- Unexpectedness, definition of, 284
- V**
- Victim blaming, 159, 162–163, 171
 - approach vs. avoidance orientation, 165–166
 - experiential vs. rationalistic processing, 166–168
 - focus on self vs. others, 164–165
- Victimization, 167, 171
- W**
- War on cancer, 341, 344
- Wegner's ironic monitoring theory, 80–81
- Weight
 - management, 319

- mindsets of body
 - cultural variations, [332](#)
 - development of, [330–331](#)
 - dieting-related products, [324](#)
 - implicit theories, [324–325](#)
 - incremental theory, [331–332](#)
 - low-cost intervention efforts, [324](#)
 - management strategies, [323](#)
 - and self-regulation, [324–327](#)
 - stigma, [327–329](#)
 - obesity. *See* Obesity
- Weight stigma, [333](#)
- Well-being, [55](#)
- Willpower. *See* Self-control