

Huff
Furchert

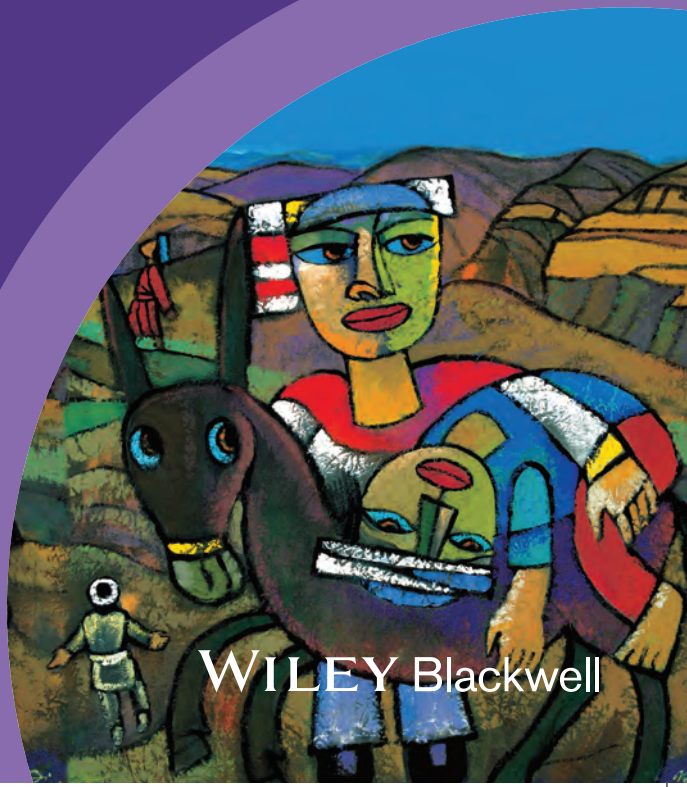


CONTEMPORARY SOCIAL ISSUES

Taking Moral Action

Chuck Huff
Almut Furchert

Taking Moral Action



WILEY Blackwell

WILEY Blackwell

Taking Moral Action offers a timely and comprehensive overview of the emerging field of moral psychology, introducing readers to one of the most vibrant areas of research in contemporary psychology. With an inclusive and interdisciplinary approach, authors Chuck Huff and Almut Furchert incorporate a wide range of empirical findings, scientific theory, and their interconnections with philosophical and theological approaches. This allows them to systematically explore the complex network of influences, contexts, and processes involved in producing and structuring moral action.

- Offers a framework of the variation in ways that individuals experience and construct moral identity and engage in moral action and formation
- Reviews cultural, organizational, group, and social influences to explore how individuals actively shape their moral environment
- Discusses the integration of reason and emotion in morality and considers how individuals can change or train their moral responses and skills
- Explores the relationships among empirical psychology, philosophy, and theology mapping out the interdisciplinary field of moral psychology
- Emphasizes the practical application of the science of morality in service of moral good

Integrating key empirical and theoretical literature, this unique volume helps readers grasp the different aspects of both habitual and intentional moral action. Thematically organized chapters examine moral action in contexts like evolution, neuroscience, and moral ecology; in influences of personality, moral identity, and moral skills; and processes of moral reason, moral emotion, and moral formation. Throughout the book, narratives of moral action and humanistic and experiential traditions of moral formation illustrate the wide variety and complexity underlying moral psychology.

Taking Moral Action is essential reading for those new to the field and experienced practitioners alike. Containing extensive references and links to further readings, Taking Moral Action is also an excellent textbook for college and university courses in areas such as psychology, ethics, theology, philosophy, anthropology, and neuroscience and in applied fields where moral action is important.

Taking Moral Action

Final Proofs

Contemporary Social Issues

Contemporary Social Issues, a book series authored by leading experts in the field, focuses on psychological inquiry and research relevant to social issues facing individuals, groups, communities, and society at large. Each volume is written for scholars, students, practitioners, and policy-makers.

Series Editor: Daniel Perlman

Multiculturalism and Diversity: A Social Psychological Perspective
Bernice Lott

The Psychological Wealth of Nations: Do Happy People Make a Happy Society?
Shigehiro Oishi

Women and Poverty: Psychology, Public Policy, and Social Justice
Heather Bullock

Taking Moral Action
Chuck Huff and Almut Furchert

Taking Moral Action

Chuck Huff and Almut Furchert

WILEY Blackwell

This edition first published 2023
© 2023 John Wiley & Sons Ltd

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Chuck Huff and Almut Furchert to be identified as the authors of this work has been asserted in accordance with law.

Registered Offices

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data

Names: Huff, Chuck, author. | Furchert, Almut, author.

Title: Taking moral action / Chuck Huff and Almut Furchert.

Description: First edition. | Hoboken, NJ : Wiley, 2023. | Includes bibliographical references and index.

Identifiers: LCCN 2023001524 (print) | LCCN 2023001525 (ebook) | ISBN 9781444335378 (paperback) | ISBN 9781118818077 (adobe pdf) | ISBN 9781118818060 (epub)

Subjects: LCSH: Evolutionary psychology. | Brain. | Social problems—Psychological aspects. | Immorality.

Classification: LCC BF698.95 .H84 2023 (print) | LCC BF698.95 (ebook) | DDC 155.7—dc23/eng/20230322

LC record available at <https://lcn.loc.gov/2023001524>

LC ebook record available at <https://lcn.loc.gov/2023001525>

Cover image:

Cover design by

Set in 9.5/11.5 Galliard by Straive, Pondicherry, India

This book is dedicated to Kelly Shaver, John Darley, and Wes Brown.

Each has been a central actor in the moral ecology that has shaped me and my approach to scholarship.

Final Proofs

Contents

<i>Preface</i>	ix
<i>Introduction</i>	xiii
Part I Contexts	1
1 Evolution	3
2 Neuroscience of Moral Action	29
3 Moral Ecology	57
Part II Influences	87
4 Personality	89
5 Moral Identity and the Self	115
6 Skills and Knowledge	145
Part III Processes	177
7 Moral Reason	179
8 Moral Emotion	215
9 Moral Formation: Shaping Moral Action	246
<i>Coda: Taking Moral Action</i>	291
<i>Index</i>	299

Final Proofs

Preface

This book is for anyone who wants to think carefully about the psychology of morality. This includes, of course, scholars and students in philosophy, psychology, and religion. But it is also for those in the many allied disciplines (e.g. criminology, pastoral care, peace studies, political science, social work, etc.) that are taking and supporting moral action in the world. More widely, we have tried to write it to be accessible to any careful reader interested in the topic, regardless of background. We hope it will provide you with a feeling for the wide variety of things one needs to consider to achieve an adequate understanding of what is today called moral psychology.

This book is an attempt to understand: (1) *why people take moral action*; and (2) *the individual's experience of actually doing so*. The first task requires us to gather many different empirical literatures on the contexts, influences, and processes that explain why people take moral action. These literatures are what has come to be called moral psychology. The second task requires us to reframe those literature reviews in terms of the experience of the individual taking moral action. This second task, then, leads us into existential, philosophical, and theological concerns. The tension between these two tasks, the general/scientific and the individual/existential, allows us to use conceptual and empirical techniques from each approach to illuminate the other, thereby helping us to understand both better.

The first author (CH) is an American empirically minded social psychologist who has pursued applied ethical issues within the world of computing and software design. One research program has been experimental investigations of gender bias in computing systems and the other has focused on life story interviews of moral exemplars in software design. His work here has been to corral the wide-ranging literatures and organize and present them in a digestible fashion.

The second author (AF) is a practicing German philosopher and psychologist with expertise in existential, phenomenological, and hermeneutic traditions as well as in the intersection of continental philosophy and applied psychology. She came to the project mainly to offer some hermeneutical tools to better map the landscape of the field, help create a narrative that connects loose ends, and bring some philosophical coherence to the categories and distinctions made in the text, thus creating the space in which the empirical focus of the text can stand.

Together, we reject the idea that it is useful to restrict moral psychology to any singular definition of morality, and instead encourage opening up the scope of “the moral” to all the interesting places one might find it. Our approach to framing the field is more

like what used to be called natural history:¹ a wide-ranging approach to collecting the phenomena of interest wherever they are. It is not theory-driven science but it embraces and uses theory-driven science. Nor is it the amateur collection of occasional specimens – we try instead to systematically seek out naturally occurring variety and pattern in the phenomena of interest. Thus, this overview of moral psychology will be broader than others. We hope this breadth helps to heal some of the fragmentation in the field, placing often-isolated literatures in conversations with each other.

Our goal is to expand the horizons of the field of moral psychology and to deepen and structure the complexities that one can find there. As you approach this book, you will find it useful to read the Introduction first, and then perhaps the short Coda. Both chapters will give you a feeling for our approach and help frame your reading of the more specialized chapters. The other chapters are designed to be read in any sequence. They are heavily interdependent, as evidenced by the numerous cross-references to other chapters within each. One way to read the book is to follow these interconnections, starting with a chapter of interest, then deciding what to read next based on your interest and the interconnections. One can also read it straight through as a broad (if not exhaustive) overview of the field.

Gratitude

A project this wide-ranging and long in gestation will have a host of people and institutions to thank. I beg your indulgence and ask that you actually read this list to give some honor to all those who have supported a project because they thought it was the right thing to do. They, of course, have no responsibility for errors or omissions in the project. I would, however, appreciate hearing from you if you find places where the text might be improved.

Thanks to Mark Snyder, the SPSSI books series editor, who read an initial book proposal that was focused narrowly on moral exemplars in computer science and helped transform it into a much more ambitious project. Thanks are also due to the initial reviewers Mark recruited who agreed that the ambitious project was a great idea but also likely impossible to complete. Fourteen years later, we can say the reviewers were right on both points.

Dan Perlman, the editor of the SPSSI series who took over from Mark Snyder, has been an unceasing voice of optimism and support in a long and difficult project. He helped us keep going when things were stalled, provided close reading and incisive commentary of his own, recruited a group of advisors and reviewers who have shaped the book in many ways, and more importantly, helped us to decide when to continue and when to quit improving things. Dan recruited several reviewers for early drafts of the manuscript. All were very helpful, but I want particularly to thank Jessica Salvatore for her close reading of almost the whole manuscript. She suggested revisions that opened new dimensions of variation in moral action and provided her encouragement to continue.

Laura Barnard Crosskey, my student and co-author on several papers, initially agreed to help as second author. While she was completing her doctoral work in Clinical Psychology at Duke (with a parallel MDiv), she wisely decided that she could

¹And is called by historians of science “Humboldtian Science” (see the Introduction for more on this).

no longer participate. Still, the project shows signs of her influence in its interdisciplinary approach and its careful attention to human detail in the case studies. She is now a clinical psychologist in Chapel Hill, NC.

Ken Fleischmann, Bill Frey, Charles Harris, Francis Harvey, Deborah Johnson, Keith Miller, Helen Nissenbaum, Michael Pritchard, Simon Rogerson, and Katie Shilton have been early colleagues in work on ethical design of software. Though none are psychologists, I have learned more from them about moral action than from many psychology textbooks. Philosophers Rosalind Hursthouse and Linda Zagzebski, both experts in virtue theory, were gracious in conversations that helped me check my understanding of the area. Philosophers Anthony Rudd and John Davenport provided the same gracious service on narrative and the self.

Both Almut and I have been hospitably received at numerous monasteries in the United States and Europe as we have taken advantage of the scholarly atmosphere in these places to write, and to re-re-write. Foremost among these is the *Studium* program run by Sister Ann Marie Biermaier of St. Benedict's Monastery in St. Joseph, MN. She and the Benedictine community there have admirably displayed the virtue of hospitality to scholars and interdisciplinary projects from across the globe. We have spent many cold (and some warm) months there shaping this book. Our daughter, now four years old, considers the monastery a second home. They have managed, without complaint, to clear almost all the crayon markings from the walls and furniture. Our daughter has also inspired the book cover, taken from a Kees de Kort illustration for the good Samaritan, her favorite bedtime story.

Other houses that have welcomed us for writing, and thereby contributed to this book, include the Jesuit Graduate School of Philosophy in Munich, specifically Professors Rüdiger Funiok and Eckhard Frick; St. Ottilien Arch Abbey in Eresing, Bavaria, particularly Fr. Otto Betler; St. John's Abbey, Collegeville MN and Stadtkloster Segen, Berlin.

Simon Rogerson at DeMontfort University, Leicester UK, provided a year's housing, the collegiality of the *Center for Computing and Social Responsibility*, and access to his extensive network of connections to support my work in moral exemplars in computing, which in the end produced the book proposal that morphed unexpectedly into this book.

The empirical and theoretical work that led to this book has been partially supported by a series of grants from the National Science Foundation [DUE-9980786, DUE-9972280, SES-0217298, and SVS-0822640]. Of course, any views expressed here do not necessarily reflect the views of the National Science Foundation. They may not even represent our current views when you read this (see the Introduction for how some theorists' definitions of the moral domain change over time). We have also received generous sabbatical and other grant support from St. Olaf College. In the liberal arts tradition, my colleagues in the computer science, philosophy, psychology, and religion departments have been very supportive, often reading drafts of chapters related to their specialty and commenting on them.

The text has benefited from over ten years of classroom testing in my advanced seminar on *The Psychology of Good and Evil*. The students' essay exams helped me understand which points were clear and which needed reconceptualization. Their comments in class shaped the creation of the ideas in revision. Like any good seminar class, we together created theory.

People who responded to queries, sometimes simply with attached PDFs, but more often with extensive conversations about work in their specialty, include: Adam Alter, Gordon Bear, Daniele Bertini, Evamaria Bohle, Nathan Cartagena, Michael Davis, Nathan DeWall, Nick Epley, Charles Ess, Jeremy Frimer, Brendan Gaesser, Richard Gaillardetz, Michael Gill, Patrick Henry, Steven Hitlin, Darrell Jodock, Matt Keefer, Scott Krepsky, Madoka Kumashiro, Thomas Lambert, Daniel Lapsley, Emmett Larsen, Elliott Lee, Jeremy Loebach, Tina Malti, Deborah Mower, Gary Muir, Michael Mumford, Danny Munoz-Hutchinson, Darcia Narvaez, Helen Nissenbaum, Monisha Pasupathi, Leigh Phillips, George Pollak, Alan Preti, Americus Reed, Sabine Roeser, Jessica Roth, Alex Rothman, Jochen Sautermeister, Douglas Schuurman, Shalom Schwartz, Anthony Seebok, Katie Shilton, June Tangney, Kris Thalhammer, Steve Thoma, Stefanie Tignor, Carlo Veltri, and David Wulff.

Among library staff, I want particularly to thank Kasia Gonnerman for her extensive interdisciplinary expertise that allowed me to find, and more importantly to access, literature in a wide variety of disciplines. She once used her research skills to help me return a wallet I found in Chicago to its rightful owner in a small village in southern Poland. After Kasia moved on to greener pastures, this project was ably supported by librarians Brian Conlan, Charlie Priori, and Audrey Gunn.

And, of course, I am deeply grateful to my wife, partner, and second author Dr. Almut Furchert. First, for continuing to support the long, fourteen-year journey of this book even as it stole precious time from us. Second, for bringing perspective and philosophical theory that has turned my fragmented intuitions about moral psychology into something resembling a philosophical framework and has shaped the psychological conceptions we propose. And finally, for her close editorial eye that has made each chapter far more readable than it might have been. She has helped me to become the scholar, and the person, I have hoped to be.

Chuck Huff
March 2023
Stadtkloster Segen, Berlin

Introduction

“Research that produces nothing but books will not suffice.”

Kurt Lewin (1946, p. 35)

“There is nothing as practical as a good theory.”

Kurt Lewin (1945, p. 129)

Why are people good (or not)? What motivates moral actors and moral action? What is moral action and which actions are moral? In ethics and morality, these are all beginners’ questions and puzzles pondered across the ages and cultures. To phrase them in psychological terms: What contexts, influences and processes, both within and outside the person, structure and support moral action? Or more briefly: How do people take moral action and become more (or less) moral?

But even these oddly phrased questions are not simply psychological. In multiple cultures, they have been studied and contemplated in many strands of philosophical inquiry and by numerous devotees of various religions. And they operate at least at two levels: an abstract and theoretical one and an urgently personal and practical one. We want both to know what the good is and how to become good.

Let us begin with this tension between the abstract/theoretical and the personal/practical. The two quotes that begin this Introduction are from one of the founders of social psychology. They seem at odds: one denigrates theory and the other praises it. But Lewin’s point was more subtle. As both a great theorist and a founder of applied social psychology, he knew from experience that abstract puzzles untethered from real social situations produce a misleading picture of the complexity of the human condition. He also knew that good theory, fittingly applied, can drive practical change in the world. Lewin’s exhortation to balance theory with practice is surely relevant in the rapidly unfolding field of moral psychology. But achieving this balance is not just a methodological issue: *it requires some grasp of the range and complexity facing the field.*

That is the task we grapple with in this text. Our attempt to provide a first overview of the emerging but highly fragmented field of moral psychology is for both those beginning in the field and those deep in the weeds and thickets of theoretical controversy. It hopes to account for the complexity of the human condition as it is lived in

real lives (Callaway & Strawn, 2020), and to provide a framework that will help harness the explanatory power of moral psychology in understanding and shaping moral action.

The title *Taking Moral Action* is an attempt to encompass the complexities of the field. Moral action explicitly names a behavioral outcome, the actual action that we call moral. The verb “taking” implies a complex web of processes accompanying the action: interior actions like planning, practicing, critiquing, judging, and evaluating that might precede, guide, and evaluate our exterior moral action like helping, organizing, protesting, or even refusing to act. Action itself signifies the intentional character of moral action (contra behavior, response, etc.), though it also allows for habitual or highly routinized performances. And it presumes a context within which that moral action makes sense, a context that both facilitates and constrains the action. Lastly the attribute “moral” signifies that the action we are interested in is related to a moral standard, value, or goal; a moral “ought” that signifies the normative expectations, evaluations, and judgments that cultures, societies, groups, and individuals make.

The chapter structure of the book is a first attempt to gather, structure, and integrate the scattered but profound empirical literatures that help us grasp different aspects of the complexity of the interior/exterior and habitual/intentional acts that we call moral action. Thus, we gather and discuss empirical literatures that consider: (1) knowledge of the *evolutionary, neurological, and cultural contexts* of the human condition; (2) the *influences* of *personality, moral identity*, and the *skills and knowledge* needed to support moral action; and (3) the *processes* of *moral reason, moral emotion*, and *moral formation* that occur both instantaneously in any situation and across the life span.

Keeping this broad span of topics in mind can help, in hermeneutic terms,¹ to expand the horizon of our understanding of the issues at hand (Gadamer, 1975; Huff & Furchert, 2014, 2016) and to avoid the pitfalls and dead ends associated with the fragmentation and isolation of the field.

I.1 A Very Brief History of Moral Psychology

Some of the earliest empirical work in moral psychology was an attempt in the 1920s to understand how moral character dispositions, such as honesty, influenced moral action (Hartshorne & May, 1928; Hartshorne, May, & Maller, 1929; Hartshorne, May, & Shuttlesworth, 1930). What emerged from this early work was a conviction that there was very little consistency. This search for consistency in moral behavior remains an issue that troubles moral psychology today. These findings discouraged further work on character for a while, and moral issues became a focus in other sub-disciplines of psychology. Kurt Lewin, one of the founders of social and organizational psychology and an importer of the German Gestalt tradition, did work on patterns of group organization (authoritarian, democratic, laissez-faire) that encouraged internal (dis)obedience and aggression (Lewin, Lippitt, & White, 1939). Work

¹The art of interpretation, understanding, and communication.

by Gordon Allport (1954/1988) established the field of prejudice research. Stanley Milgram's "shock studies" on obedience to authority made social psychology popular and controversial (Milgram, 1963, 1974). John Darley and colleagues' work on helping (Darley & Batson, 1973; Latane & Darley, 1970) spawned a still-lively cottage industry of research in pro-social behavior (Dovidio, Piliavin, Gaertner, Schroeder, & Clark, 1991; Dovidio, Piliavin, Schroeder, & Penner, 2006; Everett, Haque, & Rand, 2016; Habashi, Graziano, & Hoover, 2016; Penner, Dovidio, Piliavin, & Schroeder, 2004). All these areas of research (and others we have not named) were in part motivated by their connection to morality. Some (e.g. Milgram, Darley) have been more thoroughly integrated into current research in moral psychology than others (e.g. Allport, Lewin).

But it was Lawrence Kohlberg's groundbreaking work that constructed the first psychological home for moral psychology.² In the introduction to his dissertation Kohlberg (1958, pp. 1–3) signals his intent to challenge the reigning psychological approaches to moral development. He viewed them as relativistic, dangerously reducing moral judgment to simple cultural socialization. This socialization approach did not, in his eyes, recognize the special claim that judgments in the moral realm have – that they are about what is good and right, and that *one can transcend and rationally critique one's culture's judgments*. His concentration on the conscious act of moral judgment, and his grounding in both mainstream developmental psychology and philosophy produced a widely influential research program and tradition that had great empirical, theoretical, and applied success. It was so successful that moral psychology became a wholly owned subsidiary of developmental psychology. Social and personality psychology research was not easy to incorporate into this scheme, partly because from within the Kohlbergian paradigm of conscious, independent judgment it appeared to represent rejected forms of influence, or bias (e.g. situational factors, personality characteristics, conformity).

It was not until the collapse of the Kohlberg paradigm in the 1980s that social and personality psychology work became part of the mainstream in moral psychology again. The paradigm collapsed in part because of the consistently small correlations of moral reasoning ability with actual moral behavior.³ It was replaced with a wide range of theoretical and empirical contenders and approaches: neuroscientific, evolutionary, personality, situational influence, emotion, culture, cognitive, and others. Many of these research programs have recently begun to establish interdisciplinary collaborations with philosophy, anthropology, etc.

This flowering of different approaches is both an opportunity and an obstacle for a coherent psychology of moral action. Some researchers have claimed a "new synthesis" for this body of work (Haidt, 2007). However, for many others the jury is still out, with Frimer and Walker (2008, p. 335) claiming that moral psychology is still "pre-paradigmatic." Sinnott-Armstrong (2016) has argued that it might be impossible to assemble a single theory to unite the field (see also Greene, 2014, 2015). Ellemers, van der Toorn, Paunov, and van Leeuwen (2019) conclude their extensive

²See Chapter 7 for a short history of this.

³Both words in the phrase "consistently small" are important. See Chapters 4 and 7 for more detail on this.

review of the field by claiming that it is “unbalanced, neglect[s] some key features of human morality ... and [is] not well integrated.”⁴ Our goal in this text is not to provide a theoretical synthesis or integration – it is instead to present an explanatory, hermeneutic framework that can gather these various and partially overlapping literatures into one place where they can better inform each other.

I.2 A Fragmented Field

Moral psychology is indeed a fragmented field. To begin with, it is actually two fields: one in psychology and one in philosophy. Further, in both disciplines, there are numerous relatively isolated cottage industries of research that ignore not only the other discipline but also other literatures within their own discipline. Within these industries, the different approaches tend to be framed as adversarial and exclusive, rather than integrative (D. I. Walker & Thoma, 2017). The intense focus of these areas is often excused by the wholly true claim that a narrow focus allows them to make excellent progress. However, the doubtful claim that the narrow focus is centered on some kernel that is really representative of all morality is often implied. Superficial claims of what the specialized research tells us about “morality” restrict our view of morality to some smaller aspect of it: judgment, intuition, virtue, etc. In this way, the wide breadth of phenomena we are tempted to call moral action is implicitly reduced to some more easily studied topic. It is as though we wanted a necklace but have instead constructed a scattered and incomplete collection of pearls. There is tremendous productivity of research but no narrative thread that ties it all together. Ours is a first attempt to collect the pearls and to offer a string, or at least to frame the parameters for what such a string would look like.

A second difficulty with the field of moral psychology is its isolation from any study of real moral action. Much work in moral psychology (both psychological and philosophical) has been on the research/theory side, and shied away from actual application. This tension has been there from the beginning, as exhibited in the two earlier quotes from the founder of applied psychology, Kurt Lewin. The tension is more clearly revealed in the name for a research method Lewin pioneered, “action research” – a combination of social action and research on its effects (Adelman, 1993). We are used to thinking of research and theory as abstract knowledge and to relegating those who practice in some applied field (medicine, technology, law) to the duties of “applying” that theory. However, action research erases or at least elides this distinction, in favor of learning while doing. This is in part because Lewin thought that both theory and practical action were important and enhanced one another. This is another reason this text offers a focus on moral action – as a corrective to a field full of data and theory about “morality” in general, with only the occasional glimpse of, or reference to, particular individuals taking moral action in their daily lives.

As the philosopher Gadamer has argued, one cannot so easily split the theoretical and the practical. Gadamer (1996) shows that, in the field of medicine, the work that doctors, nurses, and other professionals do in their normal caseload includes kinds of

⁴See also the trenchant critique of the field by Colby and Damon (2015).

knowledge and expertise that are not at all referenced in the official theory – and that simple application of theory to the concrete case is not the rule-bound exercise it is often taken to be. In fact, action in the concrete case and knowledge of theory form a “hermeneutical circle” that is required to understand the whole (Gadamer, 1975; Malpas, 2018). Here, the practitioner is confronted with the problem that appropriate application requires intimate knowledge of the theory; but in order to be useful, theoretical knowledge requires intense engagement within the context of application. Gadamer’s conclusion was that theory and experience, general knowledge and intimate understanding, statistical averages and the unique case were inextricably intertwined. Neither could be complete without the other: real knowledge requires experiential immersion; theory informs practice and practice informs theory.

Thus, given what we have called the application problem inherent in all applied fields,⁵ a central claim of this book emerges: We need a robust experimental and theoretical moral psychology that is connected to and informed by the wide variety of moral action we find in the world. For this reason, we begin our text with seven historical stories of moral action and end it with three humanistic/experiential traditions of moral formation. The traditions are presented in Chapter 9, where they help us toward a more integrated perspective. We introduce the stories here to help us map the landscape of “taking moral action.”

Each chapter will end with one of the stories and ask both how the case might inform our research and how the research and theory might help to understand the case. As we examine contexts, influences, and processes, the cases will focus our attention on moral action, and help to provide a corrective in a field that is heavily experimental and theoretical. Each chapter will also highlight the applied literature that is often lost in the scuffle of abstract theory. Properly chosen, the cases will help to emphasize the wide variety of things that count as moral action in various of settings. This will in turn challenge us to better understand the underlying contexts, influences, and processes that might connect these disparate examples, and encourage us to search for even more variety in the field. Thus, a hermeneutical circle of understanding is constructed, with the goal of better theory and better application.

After introducing the stories, we wrestle with defining the field and eventually abstain in principle from any definition. We then introduce some underlying tensions in moral psychology that each chapter in the book will need to navigate. Finally, we introduce the themes and chapter structure.

1.3 Stories of Moral Action

1. Nathaniel Borenstein is a well-known computer scientist who was and remains a deeply committed pacifist. Despite this moral commitment, in 1987 he found himself in a consulting role for a NATO working group on the design of “embedded training systems” for, among other things, tactical nuclear missile launchers. A training system is “embedded” when it runs on a system that is used in the real

⁵See Huff and Furchert (2014) for how Gadamer’s insight might be used in the applied field of computing. Here, Furchert calls this “the problem of application.”

world, but it can also run in simulation mode without actually triggering any real-world consequences, such as mistakenly launching a nuclear missile. Nathaniel only reluctantly accepted a consulting relationship with NATO when he became aware that embedding training systems in certain applications was an “extremely dangerous idea” (Borenstein, 1989) because it might result in accidental missile launches. To his great surprise, he found the military members of the working group to be very interested in his critique, which led to significant changes in recommendations for the design of training systems. Still some of the academic and industry consultants were unpersuaded, since they depended on the approach for funding for research and development.

2. Rick Munson was “running fast” in his real estate business when he and his wife decided to take in some foster children.⁶ The couple ended up founding a charitable organization and a facility that provided integrated services to physically challenged children. Rick had been raised in the Southwest of the United States in a very religious household, where his mother was a social worker and his father a minister. He and his wife were regular churchgoers. After several engagements with the foster child agency, they were asked to take in two children with physical handicaps. The difficulties associated with finding care for them led the couple into increasing commitments: volunteering at schools, serving on boards of charitable organizations, etc. Their initial commitment, and its extension to the wider community, not only transformed the lives of their foster children but eventually touched a much larger group of people. The Munsons’ “involvement in community service does not seem to be a direct result of ... religious faith,” the study concluded (Hart, Atkins, & Southerland, 2006, p. 590). Instead it seems driven more by their initial empathy for the children they fostered. Religion, however, does seem to play a role in coloring and shaping the language the couple use to describe it.
3. During the 1976–1983 dictatorship in Argentina, Anna’s⁷ husband, two brothers, and pregnant sister-in-law were all disappeared by the regime.⁸ Though Anna had sole responsibility for two young sons, she and her mother began researching answers to the disappearances of their family members. They both joined in the early marches of the Mothers of the Plaza de Mayo that eventually brought worldwide attention to the brutality of the Argentinian regime and the courageous resistance of the Mothers of the Plaza. Anna also joined in the attempt to document the extent of the disappearances, and later joined a separate organization, the Grandmothers of the Plaza.⁹ This group still today uses popular campaigns and

⁶Rick Munson is a pseudonym for a participant in a study of volunteerism (Hart & Atkins, 2006). The report of the study does not give a pseudonym for his wife, though it makes it clear they made decisions together.

⁷This is a pseudonym for one of the interviewees in Thalhammer (2001).

⁸To be disappeared as an adult was to be taken without warning, then likely tortured, and almost certainly killed. There is now a legal term (forced disappearance) in international law, thanks to the advocacy of the Argentinian human rights community (K. Thalhammer, personal communication, September 6, 2019).

⁹You can still find their active campaign to track the kidnapped children at <https://abuelas.org.ar/>.

genetic testing to search for those children who were kidnapped. Their website speaks of the struggles motivated initially by the search for family members, but it also appeals to universal human rights in its outreach to others. As of 2019 (K. Thalhammer, personal communication, September 6, 2019), Anna was still searching for a missing niece who was born in captivity and thought to have been adopted by a family supportive of the military junta.

4. In a bar in Boulder Colorado, a man wearing a yarmulke (a skull cap worn by orthodox Jewish men) was attacked by two men (Beitsch, July 20, 2016). Oddly, one of the attackers was visiting the bar with a Jewish friend. When he was charged with a hate crime, he did not understand why. Boulder has a local program of “restorative justice” (Braithwaite, 2002, 2014) in which, after preparatory meetings for both sides, offenders sit down with victims and other relevant parties (in this case, the bar owner) to tell their stories and come to an agreement on how the offender can repair the harm. The agreement is often overseen by a justice official, such as a probation officer. In this case, the victim was able to say to his attacker “What if I didn’t have a yarmulke on? Would you still have assaulted me?” This brought a moment of clarity to the discussion and to the offender. As a result, the victim made it clear that he did not want the harshest sentence to be given. The judge in the case allowed the offender to serve his sentence in stages so that he could keep his job. The bar owner confessed his concern that the bar had an unsavory reputation for fights and the other assailant in the case promised to design a mural about inclusion for the bar. As in any criminal case, the motives of the defendants are colored by self-interest in avoiding punishment, but one aspect of restorative justice is helping to shape additional motivations that might lead to better outcomes.
5. Thomas Merton, born in France to an Anglican father and Quaker mother, became, by the end of his life, one of the most widely read modern Catholic authors. He was also an influential advocate for interreligious dialogue and for nonviolent struggle and human rights in America. For this latter aspect of his work, he was criticized by many for political writings “unbecoming of a monk” (Thomas Merton Center, n.d.). One can, for example, find scathing critiques of racism and “white civilization” in his *Conjectures of a Guilty Bystander* (Merton, 1966) or his *Letters to a White Liberal* (Merton, 1964). Merton’s radical political stance was an integral outgrowth of his spirituality, of his singular desire to re-form the self in search of a relationship with God (Kilcourse, 1996). At least in biographical terms, the search for God came first: “All I knew was that I wanted grace, and that I needed prayer, and that I was helpless without God” (Merton, 1948/1998, p. 329). Merton’s response to the turbulent decades of the civil rights movement might have been further isolation in the monastery. But a complex web of motivations and opportunities led him instead to the forefront of the civil rights movement in the 1960s where he became one of the most widely read Catholic authors of the century.
6. After it became clear in November of 2020 that Joe Biden had won the US presidential election, then-President Donald Trump engaged on a long campaign to invalidate the results and remain in office. Two relatively obscure figures became the early public faces of the resistance. Aaron Van Langevelde was one of two Republicans on Michigan’s Board of State Canvassers whose job it was to certify

that the votes in the election had been correctly tallied. He was under intense pressure from high-ranking party members to vote no, abstain, or resign. He voted yes because the cost of not doing so would be “constitutional chaos and the loss of our integrity” (Bidgood, 2021). Brad Raffensperger was the Secretary of State in Georgia. He received a long phone call from then-President Trump in which Trump both threatened Raffensperger with federal investigation and asked him to “find” 11,780 votes, just enough to throw the state’s result to him. Raffensperger’s reply on the phone was “We believe that we do have an accurate election” (Gardner & Firozi, 2021). In a subsequent book (Raffensperger, 2021), he explained his motivation as choosing “integrity and truth” vs. “what is expedient.”

In the Fall of 2021, US House Representative Elizabeth Cheney (Wyoming) sacrificed her high-ranking position in the Republican Party to co-chair the investigation into the January 6 attack on the US Congress. She had previously won her party’s endorsement to Congress with 73% of the vote and she “knew she could easily have done the same again” if she simply went along with her party’s widely accepted (or at least not rejected) narrative of a stolen election. “That was a path I could not and would not take,” she said in her concession speech, insisting that she had a higher loyalty to “the principles that we are all sworn to protect” (Fung, 2022).

7. Sophie Scholl has become the icon of stories of German resistance against the Nazi regime, in no small part because of her elder sister’s publication of a book about the White Rose movement in which Sophie and her brother Hans were central actors (Beuys, 2012). Her parents were, from the beginning, suspicious of the National Socialist (Nazi) movement, but the five Scholl children were enthusiastic supporters. Sophie, who had joined the League of German Girls (a branch of Hitler Youth) in 1934 did not give up her membership until 1938, and still attended “home meetings” as late as Spring 1941. Sophie’s published correspondence with Fritz Hartnagel (Hartnagel, 2005), an officer in the German army, documents her struggle to clarify her attitude toward her country and the National Socialist movement. When Hans founded the White Rose resistance movement at the University of Munich in the summer of 1942, Sophie, recently enrolled at the university, was aware and supportive of the effort. Later she was an active author of the fifth and sixth leaflets the movement produced and distributed. These were addressed to “the intellectuals” of Europe calling them to reject the regime, and they eventually found their way to Great Britain, where they were reprinted and dropped over Germany by plane. In February 1943, Sophie and Hans were leaving leaflets in stacks in the main building of the university. When she noticed there were a few left in their satchel, Sophie dug them out and tossed them over the railing of the hallway into the main atrium of the building. A janitor confronted them and turned them in to university officials, who had them arrested. After interrogation and a very short trial, they were both executed by the Nazis in private (Beuys, 2012).

All seven of these very different historical examples involve taking moral action. Some involve directly helping others, some involve working to change systems, to build community, or to advise, educate, or support others. Some, like the Munsons or Sophie Scholl, begin with small commitments that slowly balloon into larger ones, sometimes much larger. Other examples, like Nathaniel Borenstein or Thomas Merton, involve

commitments to one goal that transform themselves into other related commitments. Many stories come with a turning point, a change of heart, a break with formerly held beliefs in pursuit of a greater good. Some take place in contexts of moral conflicts where one's moral action conflicts with someone else's conviction of the right (e.g. Sophie and Hans Scholl vs. the janitor, Cheney vs. her party). Others, like the Munsons, are a deepening of commitments to a widely recognized good.

What are the goals, desires, obligations, plans, hopes, or values that motivate the very different moral actors and actions presented in these stories? What aspects of their personality, situation, and abilities help to structure their action? These are the questions that drive this book, along with a recognition of the wide variety of human action that we are tempted to call moral (or to reject as immoral). These examples actually represent six different, and mostly disconnected, literatures in applied psychology and social science:

1. Nathaniel Borenstein's story is representative of research on virtuous action in various professions. These professions include computing and engineering (Huff & Furchert, 2014; Huff & Martin, 1995; Pritchard, 1998, 2001), dentistry (Bebeau, 1994), journalism and public relations (Plaisance, 2014a, 2014b), and law (Hamilton & Monson, 2012). Much of this work was inspired by early research on "moral exemplars" in the contexts of the Holocaust (Oliner & Oliner, 1988) and of social service in America (Colby & Damon, 1992) and it was based in part on theological and philosophical virtue models of moral action.
2. The Munsons' story is drawn from a large literature in volunteerism (Stukas, Snyder, & Clary, 2016) that looks at why people volunteer in various contexts (Atkins, Hart, & Donnelly, 2005; Snyder & Omoto, 2008; Stukas et al., 2016) and how that experience of helping others shapes their self-concept, moral development, and future moral action (Grube & Piliavin, 2000; Hart et al., 2006; Hart & Matsuba, 2007; Snyder & Clary, 2004).
3. The Mothers of the Plaza and Sophie Scholl are both examples of what has been called "courageous resistance" (Thalhammer, 2001): a much more high-stakes volunteerism in which people often risk their own welfare and lives to resist actions and policies of government. Much progress has been made in constructing psychological and sociological models of how this highly risky form of moral action is nourished, motivated, and organized (Comer & Sekerka, 2018; Thalhammer et al., 2007).
4. The restorative justice story from Boulder gives us some insight into how people might move from being below acceptable standards of behavior to becoming valued members of the community. Recent movements in restorative justice have begun to focus attention on the psychological processes by which criminal offenders move from what are called "criminogenic beliefs" that excuse their offenses to a more objective understanding of how society views them (Braithwaite, 1989, 2002, 2014, 2016; N. Harris, 2006; Tangney, Mashek, & Stuewig, 2007). The transformation of the moral emotions such as shame and guilt has been highlighted as a central process in this rehabilitative change (Stuewig, Tangney, Heigel, Harty, & McCloskey, 2010; Stuewig, Tangney, Mashek, Forkner, & Dearing, 2009; Tangney, Mashek, et al., 2007; Tangney, Stuewig, & Hafez, 2011; Tangney, Stuewig, & Martinez, 2014; Tangney, Stuewig, & Mashek, 2007). How we think about guilt, blame, and moral becoming may undermine or support this rehabilitative

purpose (Darley, 2001; Darley, Carlsmith, & Robinson, 2000; Gill & Cerce, 2017; Gill & Getty, 2016).

5. The story of Thomas Merton is one being drawn by “ultimate verities” (Luhmann & Weisman, 2022; Maslow, 1969) that naturally developed to include moral commitments. One of the serendipitous findings of the classic work in moral exemplars (Colby & Damon, 1992) was that even though the selection criteria did not include reference to religion, almost all of the exemplars saw their moral commitments as outgrowths of, and supported by, their religious commitments.¹⁰ Citing this finding, Walker and Frimer (2008) begin their article on “Transcendence in the lives of moral heroes” with the claim “Moral psychology suffers from a grievous blind spot” and go on to detail the complex and difficult relationship of transcendence and morality. They argue, as we do in this book, that there are multiple paths to mature moral functioning and many include religion. Religion’s relationship with morality is fraught and paradoxical, promoting both moral and immoral action (Bloom, 2012; DeWall et al., 2010; Ginges, Hansen, & Norenzayan, 2009; Graham & Haidt, 2010, 2012; Koole, McCullough, Kuhl, & Roelofsma, 2010; Ysseldyk, Matheson, & Anisman, 2010). As we will discover in Chapter 9, this may not be an effect of religion per se, but more a reflection of the moral collapse (Hart, 2005) that can happen when a relative value becomes falsely sacralized (Graham & Haidt, 2012) or is taken as being of ultimate value itself (Furchert, 2022). The pull of the ultimate plays a crucial role in all the narratives. Abraham Maslow’s often neglected final stage in the hierarchy of needs is that of self-transcendence in service of a goal higher than the self (Koltko-Rivera, 2006). These goals can be religiously inspired but also moral, artistic, social, or intellectual. All share a motivational pull of “ultimate verities” such as “truth, goodness, beauty, perfection, excellence, simplicity, elegance” (Maslow, 1969, p. 4). In this vein, we might also include exemplars like Emily Dickinson, who surely was motivated by devotion to the art of poetry to “explore the full range of human experience” and whose belief was “nimble” rather than religious (Lundin, 2004, p. 4).
6. The stories of Van Langvelde, Raffensperger, and Cheney seem at first glance to be the clearest example of people acting on principle to “do the right thing,” but there is much complexity hiding behind them. They are clearly all referring to standards of conduct that prescribe and proscribe certain actions (C. E. Harris, 2008; Janoff-Bulman, Sheikh, & Hepp, 2009). Van Langvelde and Raffensperger, however, also refer to their roles as being “public servants” who should administer their duties without partisanship. Cheney speaks of duties to higher principles than party loyalty (Karl, 2022). This suggests that at least their professional identities (Bebeau & Monson, 2012; Bebeau & Thoma, 2012) were part of what was at stake in their decisions. And the inclusion of “constitutional chaos” in Van Langvelde’s comments suggest some calculation of the costs for his party, the government, and the people. There is a long literature of how various motivations are traded off in ethical decision-making (Fleischmann, Lammers,

¹⁰This finding of extensive religious inspiration for moral action was not replicated in a northern European sample of moral exemplars in computing and engineering (Huff & Barnard, 2009).

Conway, & Galinsky, 2017; Gamez-Djokic & Molden, 2016; Shenhav & Greene, 2014). The approach that seems most fitting here is Hirschman's (1970) exit, voice, loyalty decision tree and the various influences that lead to each decision (or lack of decision that amounts to choice). The intense pressure on them all has been regularly attributed to tribalism in politics (Clark, Liu, Winegard, & Ditto, 2019) and its emphasis on loyalty (Frimer, Gaucher, & Schaefer, 2014). The complexities of multiple loyalties, and their interactions with group identity and moral principle, make loyalty a "double-edged sword" (Berry, Lewis, & Sowden, 2021) that can result in moral or immoral behavior depending on the actor's judgment of "what does a situation like this mean for a person like me?" (Weber, Kopelman, & Messick, 2004). Thus, even in what seems like a clear case of principled action, there is much complexity to be discovered.

Though anecdotal and incomplete these examples and accompanying literatures hint at the range and depth of moral action that this book tries to investigate. Moral action pervades life across cultures and has been central to human culture since we first began keeping records.¹¹ Moral traits are central to our sense of self and identity (Strohmingner & Nichols, 2014) and people report being willing to go to great lengths to protect their moral reputation (Vonash, Reynolds, Winegard, & Baumeister, 2017). Moral dissimilarity produces a desire to distance ourselves from others, intolerance of those others, and an inability to design solutions to resolve disagreements (Skitka, Bauman, & Sargis, 2005). And moral contagion drives much of what becomes "viral" on the internet (Brady, Crockett, & Van Bavel, 2020).

The seven examples we provide of real-world, applied moral action give some idea of the variation in things we include under the description of moral action, the wide range of domains in which moral commitments matter, and the variation in commitments that one can find in those who regularly take moral action. We will return to them throughout this book. Examples of other areas and dimensions of variation of moral action will accumulate as we read through the text.

I.4 Considering the Scope of Moral Action

Psychological moral psychology,¹² the field this book attempts to introduce and to organize, is one of the most vibrant and active areas of research in psychology today (Ellemers et al., 2019; Greene, 2015; Haidt & Kesebir, 2010; Huff & Gaasedelen, 2011). In addition to generating a great deal of research, it also plays a

¹¹Chapter 3 begins with an example of friendship taken from the Epic of Gilgamesh, recorded around 4,000 years ago. For an encyclopedic overview of early thinking about morality see Bellah (2011).

¹²There is a connected field of *philosophical* moral psychology, with significant overlap in topics and even some overlap in method. See Tiberius (2014) for a thoughtful and systematic overview. This field is primarily analytic in its approach and we provide occasional references to it. In contrast, the philosophical view we bring to our text is a continental/existential one that we think helps to highlight important issues in moral action and moral formation.

central role in psychology, touching almost every subfield. We provide here an introductory overview of the breath of the field without claiming it to be exhaustive.

With an inclusive approach, we hope to produce and support an *interdisciplinary* conversation by providing doors to the wide range of psychological research in moral psychology and by suggesting some narratives upon which those doors might open (Heiphetz & Cushman, 2021). Two philosophers who deal in moral psychology have written, “in science there is seldom, or never, a single decisive experiment or, for that matter, a decisive experimental failure. General conclusions about social science can legitimately be drawn only from encountering, in full detail, a body of research, and adducing patterns or trends” (Doris & Stich, 2005, p. 123).¹³ We hope that by adducing patterns in psychology we can also point out doors that open upon both philosophy and theology. Thoughtful readers will be able to make connections to fields even further removed.

1.4.1 Exploration Instead of Definition

To capture moral action in all its variety, we must resist drawing narrow definitional lines that would exclude morally interesting phenomena. To be perfectly clear, *we reject the idea that it is useful to restrict moral psychology to any singular definition of morality, and instead encourage opening up the scope of “the moral” to all the interesting places one might find it.* Thus, we open the focus of the text to the production and maintenance of moral *action* (rather than concentrating, e.g. on moral judgment or intuition¹⁴). We include making judgments but also the experience of moral emotion, and the practice of sharpening those emotions. We include the influence of organizations on moral action and the influence of moral action on organizations. At the concrete level represented by the stories with which we began, we want to include taking action to change social systems, regularly caring for a foster child, writing guidelines for military weapons systems to make them safer, seeking an authentic relationship with one’s spouse or with God, a singular devotion to plumbing the human soul through poetry, trying to become a law-abiding citizen or make restitution for a crime, looking for your stolen grandchild, being loyal or disobedient, and a host of other interesting and even contrary things people do that one might consider at least in part morally motivated.¹⁵

How does our scope differ from others that researchers in the area have used? The Selective Table of Definitions of Moral Psychology at the end of this chapter provides sixteen different definitions of morality that have been used in the field. Many are still

¹³We would add that this applies to all science, from physics to medicine and technology (Collins & Pinch, 2006).

¹⁴Still, one can find wide variation even if one limits inquiry to moral judgment. Malle (2021, p. 294) lists four different types of moral judgment that have been often confused in the literature (evaluations, norm judgments, moral wrongness judgments, and blame judgments) that all differ in their typical objects of judgment, the information they process, and their social functions.

¹⁵In terms of philosophy of science, this attempt to shift the focus is called “tree switching” (Thagard, 2014). It involves building a different structure of the phenomena of interest that allows one to think in new ways about the area.

in circulation. Some come from the same theorist whose ideas have changed over time or context (e.g. Haidt, Kohlberg, Narvaez). One might find that both a quick read of this table and longer study of additional definitions will reveal there is no theoretical unity in definition or area of focus. For this reason, we have included a column that gives some idea of the purpose for which each definition was crafted.

Some of these definitions of morality are quite restrictive (e.g. 1 and 9, focusing on judgment), while others are quite open. For the purpose of this text, our preference is to embrace some synthesis of the most expansive ones (e.g. 4, 10, and 14) and simultaneously to reject imposing definitional restrictions. This allows us to search widely for interesting moral phenomena.

We liken our broad approach to that of nineteenth-century natural history. Historians of philosophy have called this approach Humboldtian Science in a reference to the great German scientific explorer, Alexander von Humboldt (Cannon, 1978; Dettelbach, 2001). Charles Darwin was inspired by Humboldt's scientific travels, took Humboldt's works on his famous voyage on the *Beagle*, and modeled some of his writings after Humboldt (van Wyhe, 2002). At its best, natural history was a wide-ranging approach to collecting the phenomena of interest wherever they were, a careful attention to measurement and concern for error, and an openness to looking for connections by immersing oneself in the phenomena (Cannon, 1978). It was not theory-driven science, but it embraced and used theory-driven science. Neither was it the amateur collection of occasional specimens, but it systematically sought out naturally occurring variety and pattern in the phenomena of interest. It cherished controlled experimentation, but also careful observation and measurement, and intense engagement with the emerging patterns.¹⁶ We do not lay claim to the comprehensiveness of von Humboldt, Darwin, and other nineteenth-century naturalists, but in their spirit, we try to open wide the doors to the variety of moral action.¹⁷

In this spirit, we might then adopt definition 14, offered by the philosopher Sinnott-Armstrong, precisely because of its broad scope and vague application. It calls on popular “exemplars or paradigms of moral judgments” or what Lapsley and Lasky (2001) call prototypes. But even this seems to leave out moral action that does not recruit explicit judgment, possibly excluding important phenomena, such as expert or habitual moral action.¹⁸ So, we offer the table itself as a sampling of definitions that helps to specify the family resemblance of what has been studied as moral. But we leave our inquiry open to the moral wherever we might find it.

¹⁶The ideal of an attitude of eager openness to the phenomenon in its complexity and detail is still current in science, though it is not often talked about in research methods texts. Barbara McClintock, the Nobel prize-winning geneticist, describes her method of careful experimentation and close observation as depending upon a “feeling for the organism” (Keller, 1983, p. 198). The philosopher Shannon Vallor (2014) calls this scientific virtue “perceptual responsiveness” and provides an analysis and examples of it.

¹⁷One can find a parallel call for an expansion of a field in a recent critique of the emotion literature, where Barrett, Adolphs, Marsella, Martinez, and Pollak (2019, p. 50) call for research that “attempts to study facial movements in real life, measuring context, sampling across cultures (even though these studies are often less well controlled than studies in the laboratory).”

¹⁸See Chapters 6 and 7 for these approaches.

Interestingly, all definitions in the table involve implicit or explicit recognition of an “ought” that is not merely instrumental to some other goal. Some anthropologists focus on this independent and objective feeling of the “ought,” (e.g. Keane (2015) in definition 10, or Shweder and Menon (2014)).¹⁹ They describe a feeling, across all cultures, that moral judgments are claims about an objective moral reality with which other reasonable minds must agree. All of the listed definitions share something of this “ought.” However, some of the definitions, and most certainly our examples, include motivations that do not have an explicit ought in them. Merton’s search for God and social justice was not undertaken because he ought to do it but because he realized his desperate need for it. The Munsons took in foster children and engaged in social service not because they had an obligation but from other, more complex motivations; as did the Mothers of the Plaza de Mayo and the Republican officials who rejected Trump’s proposed voter fraud. So, there is more than “ought” in the psychology of producing moral action.²⁰ Still, in every case, one finds the draw of the ultimate at least at a remove from or as one among many motivations. And struggling with the question of what one “ought” or ought not do is itself a kind of moral action, an internal process that is motivated and constrained by a host of influences, some more selfish or exalted than others. This struggle can bring us back to the domain of the ultimate, or as the theologian Tillich (1951, 1952) has it, to what concerns us ultimately. Moral exemplars have often described this as feeling pulled toward a greater good, a desire to do “the right thing.” In the case of Sophie Scholl, for instance, the “ought” changed from a loyalty to Nazi philosophy to giving her life for speaking out about its cruelty.

In this quest to resolve definitively (and definitionally) the “what is moral?” question, we need to take seriously the claim of the philosopher Sinnott-Armstrong (2016) that *there may well be no unity in what we call moral*, and thus no single answer to this question. He finds, for instance, no unity in terms of specific *pathways or structures in the brain* that are always involved in moral judgment. Nor does he find unity in the *content* of morality. Is it about harm? Yes, but some morality is also about sacredness, or loyalty, or authority, or compassion. This list itself is open to revision.²¹ Nor does Sinnott-Armstrong find unity at the level of the *function* that morality serves. A functional approach involves citing “some widespread problem that moral judgments all help to solve” (Sinnott-Armstrong, 2016, p. 342). One proposal is that morality functions as a “guide to social conduct.” But this includes some inappropriate social things (e.g. local customs or fashions) and inappropriately excludes some solitary activities (e.g. some aspects of personal piety or some instances of suicide that still seem moral). Another approach is to assign to morality the function of “regulating self-interest or making cooperation possible” (Graham et al., 2011; Greene, 2015).

¹⁹Philosophers as divergent as Kant (1785/2011) and Kierkegaard (1849/2004) have also depended on it.

²⁰Especially if this “ought” is only conceived of as a Freudian superego or mere socialization.

²¹See Chapter 7 for several frameworks with different lists. One solution in the search for a unitary morality involves an attempt to derive all the other items in the various lists from a single item that is claimed to be foundational (e.g. harm or compassion), but Sinnott-Armstrong (2016) shows the difficulties of these attempts. See Schein and Gray (2015) for an empirical attempt to construct a unitary basis.

But this, too, falls short. There are many aspects of morality that are not needed to make cooperation possible, and not all aspects of morality are about regulating self-interest (indeed some aspects positively *encourage* the expression of self-interest – see Chapter 7, Section 7.3.3.4).

We conclude that it is very difficult to find a definitional fence that will appropriately exclude all nonmoral judgments and include all moral judgments (never mind extending this to actions). Sinnott-Armstrong (2016) proposes that what is needed is a taxonomy of the different kinds of things that we call moral, allowing us to study the different functions, contents, processes, and influences associated with each kind. Our preference therefore is to adopt the natural history approach we introduced earlier, and to see this method as a necessary preliminary step in producing the taxonomy.

1.4.2 Shared Horizons: Philosophy and Theology

In trying to resolve the issue of how to define the field, we have needed from time to time to venture outside so called empirical moral psychology and into philosophy and theology. Throughout this text we regularly add brief excursions into the humanities to help us grapple with certain concepts, create new frameworks or open doors for discourse between these fields. (We have tried to mark these switches so that you can be aware of them.) Such excursions are meant to mark where we believe a discourse between the fields is needed, but not to develop the discourse in any depth ourselves.

1.4.2.1 Psychological Moral Psychology As we have pointed out, psychological moral psychology is still an emerging field. Our attempt to cover the empirical and theoretical literature relevant to it requires the creation of a theoretical framework or narrative that makes its overall structure somewhat clearer. The chapter structure of the book reflects that theoretical framework. To maintain the focus on the framework, we will mostly avoid blow-by-blow descriptions of particular studies, graphs of interactions and meta-analyses, and much of the sometime internecine debates about methods and measurement. In some cases, when there are many alternative conceptions (e.g. in Chapter 9) we simply choose one conception we think best fits the narrative we are developing, and provide notice of alternatives in a footnote. Interested readers are welcome to explore the implications of using the alternative conceptions. In other cases, when we think one approach has crowded out other useful approaches we provide a short overview of the main alternatives (e.g. in Chapter 7, where we present moral modules, moral foundations, human values, and relational models as alternative approaches). We hope this approach will bring some unity to the field of psychological moral psychology and help psychological, philosophical, and theological readers to see some larger patterns and become aware of alternatives.

1.4.2.2 Philosophical Moral Psychology Moral psychology as the quest for doing good has been a philosophical problem since its beginnings. In the last decades there has been increased interest in the interaction between classical moral philosophy and empirical moral psychology. Tiberius (2014) and Miller (2021) provide an excellent overview of these well-developed literatures in moral philosophy.

Readers will find occasional reference in this text to philosophical concepts and puzzles such as free will vs. determinism, altruism vs. egoism, rationalism vs. emotivism,

and the relationship between empirical statements of “is” and moral statements of “ought.” We do not attempt to solve, or even sufficiently discuss, these hoary and resilient puzzles in ethics and meta-ethics. We do occasionally address them directly (e.g. the relation between reason and emotion) but mostly just mention them to suggest interesting connections or implications for psychological research and theory (e.g. you can find altruism in Chapter 1 and *is vs. ought* in Chapter 7).

To these usual suspects in philosophical moral psychology we also add, where appropriate, references to European traditions such as phenomenology, hermeneutics, or an existential perspective of the human condition. In this Introduction, for instance, we refer to what we call the *existential turn*, a turn away from empirical reductionism toward lived experience, and from nomothetic claims toward ideographic questions. In Chapter 9, you will find connections to the existential revival of Socratic maieutic as well as to other traditional approaches on moral formation and education. In Chapter 6, we suggest a fruitful conversation might be had with Gadamer’s (1996) understanding of “the application problem” (Huff & Furchert, 2014) and provide a way of reflecting on ethical expertise. Because this text focuses on the literature in psychology, in every case, we provide mere sketches of the possible connections rather than in-depth analysis.

1.4.2.3 Theological Moral Psychology Theologians have themselves begun to discover the riches and complexities in empirical psychology and their implications for, moral theology (Callaway & Strawn, 2020). We have already mentioned that the relationship between religion and morality is one fraught with complexity and danger. We note in Chapters 3 and 7 that religion can serve as a potent cultural resource to establish a critique of the culture, can in turn be critiqued, and even be used to critique itself. In Chapters 5 and 9 we note that monastic traditions in all major religions view certain aspects of the self as standing in the way of moral formation, an approach quite different from mainstream psychological ones that presume the centrality of self-enhancement. And when covering aspects of self-regulation in Chapter 6, we connect it to the newly emerging dialogue between psychology of religion and self-regulation.

1.4.2.4 Virtue Theories and Moral Exemplars This text was first conceived as part of an educationally motivated applied research program on moral exemplars in computing (Huff & Furchert, 2014). The empirical work gathered life-story interviews from computing professionals known for extraordinary moral commitment in their profession (Huff & Barnard, 2009). The theoretical project was an attempt to use the exemplars’ narratives and the resources of psychology to characterize the influences on moral action in the profession (Huff, Barnard, & Frey, 2008a, 2008b). Approaches that begin with moral exemplars assume that there is something about moral character that is central to the production of moral action (Damon & Colby, 2013; L. J. Walker & Hennig, 2004; Zagzebski, 2013, 2017). However, we quickly found out from our interviewees that things were more complicated. Exemplars with different personalities seemed to choose different kinds of moral projects and to approach them in different ways. Their moral careers were shaped by the organizations they inhabited, and they in turn shaped those organizations. They

seemed to spend little time in moral deliberation, rather, they concentrated on moral action. They were both deeply passionate about, and shrewdly calculating in, their moral projects. The attempt to systematize and grasp these influences and ways of being moral did not result in a virtue theory for moral action. It produced instead a moral landscape within which one could build a variety of virtue models of moral action. Or, one might say, it produced an enumeration and unfolding of the parameters for which one would need to account in any virtue model.

There are, of course, many different kinds of philosophical virtue theories (Hursthouse & Pettigrove, 2016) and all of them, with more or less modification, are likely to be compatible with this framework. Most religions also endorse the development of some set of virtues, though the sets differ both within and across religions (Augustine, 1986; Bucar, 2018; MacKenzie, 2017; Rambachan, 2006; Shanley, 2008). And psychologists have begun to construct suitably complex forms of virtue theories (Fowers, Carroll, Leonhardt, & Cokelet, 2020; Hulse & Hampson, 2014). We are not proposing or endorsing a specific virtue theory here. But we agree with the assumption of most virtue theories that individuals can develop a propensity to engage in successful moral action, a process we consider in Chapter 9. Our claim is that understanding this propensity requires at least the complexity of the framework we propose, rather than a simple reliance on internal dispositions.

1.5 Underlying Tensions in the Field

Having given an idea of the scope of our inquiry, we now turn to some underlying tensions that have on occasion preoccupied researchers in the area. This is some destructive groundwork that rejects several of the simple dichotomies that have structured debate and research in moral psychology. These debates have been tremendously productive of theory and research and have helped frame the strengths and limiting conditions of each side in each debate. However, they have occasionally been a distraction from the task of seeing the larger picture of moral action.²²

1.5.1 Personality vs. Situation

This is one of the most long-standing debates in the literature, reaching back to early work by Hartshorne and May (1928) but still being played out in current philosophical discussions of virtue and character (Doris, 2002; Doris & Stich, 2005; Snow, 2010; Stich, Doris, & Roedder, 2010). The simplest model of moral character in personality is that of single-variable influence. A child has some degree of the characteristic of honesty and the more honest that child is, the more likely he or she is to be honest in all situations. The child's character determines his or her behavior rather than any situational pressure. At least some of the children who are honest in one situation

²²We leave out of this list one significant tension: whether (im)moral action is fundamentally motivated by selfish or by altruistic goals. Because one needs some technical background to understand the discussion, we treat this at length in Chapter 1.

should be found to be honest in many other situations. Much to the puzzlement of character theorists, Hartshorne and May (1928) found it to be false. Among the children they tested, honesty in one situation had little to no predictive value for honesty in dissimilar situations. This pattern (or lack of pattern) held for many other supposedly character-like attributes they tested. In 1968, Mischel reviewed the cumulative failure of psychologists to find cross-situational consistency for personality characteristics and famously proposed that there were no simple personality traits that were broadly predictive of behavior. Thus, it seemed for a while that the situational side of the dichotomy had won the competition.²³ The review of this issue in Chapter 4 reports that the failure was more one of conceptualization and measurement. Conceptualizing a personality trait of honesty requires knowing something both about how a person views themselves and how they view the situation. Does personality or situation control moral behavior? Yes. Both do.

I.5.2 Nature vs. Nurture

There is a long tradition in psychological theory of looking for genetically determined physiological underpinnings for personality differences, including differences in morality. This is at least as old as the Greek humoral physiology and psychology (Zuckerman, 1995), though modern versions loosely tie clusters of personality traits to neurotransmitter systems and neural structure (DeYoung & Gray, 2009; DeYoung et al., 2010). But again, the answer to whether nature or nurture is in control is “yes.” Taylor et al. (2000) provide a good example of the complex interactions among nature and nurture that shape maternal helping.

I.5.3 Stability vs. Development

The nature–nurture debate is linked with the debate over whether moral character is fixed or develops over time, particularly over whether it can develop during adulthood. Those who prefer to see it as a fixed characteristic (McCrae & Costa, 1997) ground much of its stability in its physiological underpinnings. But those who argue for the developmental possibilities have conducted research programs to document adult change over time (Roberts & Mroczek, 2008; Wrzus & Roberts, 2017) and see character as at least somewhat malleable with changing life circumstances. So again, the most reasonable answer is to look for some way for both stability and development to be influential.

I.5.4 Reason vs. Bias

Excluding bias from the process of moral decision was a central concern of the classic Kohlbergian conception of morality (L. J. Walker, 2004). In this approach, bias is defined as any influences that do not conform to the ideal of “conscious rational application of general rules and principles” (see definition 2 in the Selective Table of

²³See Penner and Orum (2010) and Mischel (2004) for short analyses of the history of this controversy.

Definitions of Moral Psychology).²⁴ All influences or processes that are not based in principled, conscious, rational thought are suspicious, and therefore are thought to contain bias.²⁵ This psychological conception comes in part from Kohlberg's (1958) rejection of socialization models of morality, based in his concern that these greatly increased the dangers of indoctrination in moral education. This conception of reason and bias is at the heart of the struggle in moral psychology over whether explicit, controlled, cognition or implicit, intuitive evaluations are better descriptions of moral action.²⁶ But this produces a false dichotomy. As we progress through the review in this book, we will see that personality, moral commitment, environmental constraints, emotion, and many other items are all essential parts of moral acting, sometimes serving as biasing agents, other times as moral correctives (Huff & Furchert, 2016). We will also find that both bias and helpful correction can come from conscious thinking and from automatic processing. One cannot conceptually abstract out the "rational" or "conscious" part and declare it the sole system of measurement (Schwartz, 2010). Moving beyond these productive but eventually misleading dichotomies will require a more complex picture of the moral personality.

1.6 Mapping the Landscape of Moral Action: The Book's Structure

Having opened up this review of moral psychology to a broad range of literatures, we now need a framework to organize them. As we have said, the idea is not to attempt a unifying theory, since there is some doubt about whether a unifying theory can be constructed (Sinnott-Armstrong, 2016). It is instead an attempt to try to gather together those pearls that any theory would need to unify. One can find predecessors in this attempt: in an early departure from Kohlbergian orthodoxy, Narvaez and Rest (1995) proposed a model consisting of four processes: moral sensitivity, moral judgment, moral motivation, and implementation. Hart (2005) presented a systematic model of moral identity formation that contained many of the aspects we include here, and Lapsley (2008) provided a useful critique of it. Our framework is influenced by these proposals but also by our own wide-ranging search for relevant literatures.

²⁴ It is a basic idea of ethics that all ethical systems (and meta-ethical approaches) depend on a theory of what is good, a value theory. Philosophers and theologians have argued for millennia over what the appropriate values should be against which one might measure bias. Kohlberg chose justice as the galvanizing principle of his system (Gibbs et al., 2007; Narvaez, 2005). Others have chosen compassion (Narvaez, 2016), and others have searched for an empirical description of the values people actually use in moral life (Ciecuch, Schwartz, & Davidov, 2015; Graham et al., 2011; L. J. Walker & Frimer, 2015). Though it is not a primary purpose of our text, we will evaluate all of these options.

²⁵ In philosophical terms, this is the argument about *autonomous* (independent, reason-driven) vs. *heteronomous* (any influence other than reason) ethics. See Huff and Furchert (2016) for an empirical evaluation of the claim of autonomy. See the philosophical papers in the accompanying issue of *Dialegetha* for philosophical evaluations of this claim.

²⁶ See Chapter 7 and Frimer and Walker (2008, pp. 338–340) for evaluations of this issue.

Thus, we incorporate scholarly traditions, theories, empirical findings, and practical moral writing (all of which we call narratives) from an interdisciplinary group of sources. Given the space, we cannot provide extensive detail about each of these sources, and we often skim over or omit controversies that are contained within them. In the metaphor we used earlier, we provide doors into other narratives, and those hoping to profit from them need to look for themselves inside the narratives and literatures.

I.6.1 Overview of the Landscape

Without a definition, how do we explore the moral landscape in the way Humboldtian explorers might have done? Instead of a definition, we attempt to expand the moral landscape of phenomena of interest in moral psychology in two ways. First, inspired by an existential turn we broaden the scope of moral psychology toward the idiographic experience of the single individual as they face their capabilities, limitations, and moral goals, recognizing the wide varieties of moral goals that have a normative pull to them. Second, we offer a framework of contexts (evolution, neuroscience, and moral ecology), influences (personality, moral identity and the self, and skills and knowledge), and processes (moral reason, moral emotion, and moral formation) that help to detail the capabilities, limitations, and moral goals that motivate, empower, and constrain that single individual. Together these two innovations expand the field of moral psychology in the Humboldtian spirit by revealing both variety and structure within that variety.

I.6.1.1 The Existential Turn The psychiatrist and philosopher Viktor Frankl once claimed that there is a need for a twofold corrective in the Psychological Sciences (Frankl, 2018; Furchert, 2012). In his famous *Man's Search for Meaning* he addresses people's attraction toward ultimate values that give meaning and motivation to our lives and stories (Frankl, 1959). His complaint was that by reducing human experience to "nothing but" psychological drives, cognitive networks, or emotional appraisals, science was making a twofold mistake: First, though we experience these values as objective and universal, science has *subjectified* them into "nothing but" personal goals, and thus missed the richness of their role in our lives. Secondly, by *objectivizing* the world of subjective human experience, science has become blind to the individual's insistent longing and attraction to the world of logos and meaning making (Frankl, 1959, 1997). Thus, we study our "subjects" as objects instead of as intentional beings who are longing for and are able to self-transcend (Frankl, 1997), pulled by "ultimate verities" (Maslow, 1969), or that which "concerns us ultimately" (Hoffman & Ellis, 2019; Tillich, 1951). For Frankl it is precisely this ability of self-transcendence that holds the key to understanding meaning making, intentionality, and thus moral action.

According to Frankl, the corrective therefore needs to be twofold: (1) to bring back the world of logos based in ultimate values²⁷ (and thus the normative) into

²⁷ Maslow (1969) treats "ultimate verities," such as "truth, goodness, beauty, perfection, excellence, simplicity, elegance," as a near-universal aspect of the moral experience. It is in the service of these values (or projects that are driven by these values) that individuals transcend mere needs and self-interest. We will follow a line of research based on self-transcendence in Chapter 9 (L. J. Walker & Frimer, 2015).

psychology and psychotherapy without *subjectifying* it (as, e.g. relativism does); and (2) to find proper ways to observe the individual process of meaning making and value formation, without *objectifying* it.²⁸ Thus, to grasp the pull these kinds of values exert on individual experience, we need to open our moral landscape to them. A scientific understanding of the experience of being pulled by, or transcending self-interest, in service of a moral good calls for an ideographic turn that recognizes the normative in the existential without falling into the trap of reductionism. We are not doing meta-ethics or making a claim about whether these values are real or true. We are instead adopting a claim made in multiple disciplines about the phenomenological experience of human nature in the realm of the moral, and with it an insistence on the importance for moral psychology to understand its pull in the idiographic experience of the individual.²⁹

Thus, we argue that: (1) moral psychology should be open to the wide range of ultimate values for which individuals transcend self-interest; and (2) moral psychology is best understood through the idiographic lens of the individual's moral action and experience in all its variety. This perspective allows us to study the *normative in action*, or the *moral in becoming*, without the need to solve metaphysical riddles.

1.6.1.2 Psychological Themes How can we take inspiration from this existential turn as we immerse ourselves in multiple empirical literatures in each chapter ahead of us? To do so, we have abstracted four empirically phrased themes from the existential turn we described earlier. We think these themes provide a corrective for an emerging field that tends toward fragmentation (Ellemers et al., 2019) and emphasize the practical application of the science of morality in service to those individuals who want to become good. The themes are:

- *A broad search for variety in moral action, influences, and processes.* Research in moral psychology tends to be siloed in its conception of the phenomenon. For instance, much work on moral judgment admits the painstaking difficulty of implementation of judgment by the individual, but it omits to do research on that

²⁸One can find this double corrective emerging from the experience of normative values in anthropological work we review in Chapter 7 (Shweder, 2012; Shweder & Menon, 2014). The existential corrective is also implicit in Gadamer's claim that the theoretical and practical are bound together in mutually supportive dialogue (Gadamer, 1975, 1996). The German philosopher/psychiatrist Karl Jaspers also spans the gap between the theoretical and the practical with his claim that one must objectively reflect on general psychopathology but also subjectively consider the experience of the individual (Jaspers, 1938, 1958, 1973; C. Walker, 1995). Jaspers is partly inspired here by Dilthey's distinction between *Erklären* and *Verstehen* (Harrington, 2000). Frankl, Jaspers, and Gadamer have all been inspired by the existential corrective of the Danish theologian Søren Kierkegaard (1813–1855) and his turn toward the existential that bridges the abstract with the concrete, the theoretical with the practical, and the moral with action.

²⁹This turn is not unprecedented in empirical psychology. In some cases, it is already implicit in many models we review in Chapter 4 (Cervone, 2021; Mischel & Morf, 2012; Wrzus & Roberts, 2017). It is also explicit in some of the work we review in Chapter 6 (Kuhl, Quirin, & Koole, 2021) and in Chapter 9 (L. J. Walker & Frimer, 2015).

implementation, or to cite those who are doing that research. As a corrective, in each chapter we attempt to cast our net widely to capture variation in the topic.

- *Concentration on the idiographic experience, choice, and development of the individual.* Most work in moral psychology looks for differences in groups of people depending on some measured or manipulated variable. Work on moral identity, for instance, often contrasts a group of people high in moral identity with a group that is lower.³⁰ This is useful in seeing the influence of various factors on how people in general behave.³¹ But we are also interested in the ways that individuals experience moral identity and how they individually construct that identity. This is the idiographic level that we seek to highlight in each chapter.
- *A focus on the ultimate normative status of moral claims.* Anthropologists have found that there is broad cultural agreement that morality has a special status in contrast to other judgments. In summarizing this literature, Shweder (2012, p. 94) states that, across cultures, “moral judgments are experienced as though they are judgments about the true nature of some posited objective moral charter.” We seek to pay homage to this universal experience of special status by resisting simplistic relativism³² while at the same time recognizing the variety of values that individuals and cultures attempt to balance. We offer no philosophical or theological solutions here, but we do attempt in each chapter at least to refrain from the casual embrace of relativism in our descriptions of moral action.
- *A hope toward educational implications.* In the tradition of Lawrence Kohlberg (1958) and in the desire to support the individual becoming we describe in the existential turn, we hope the work reviewed in each chapter can be used by creative minds to structure education (for both adults and children) that facilitates and encourages taking moral action. We occasionally make suggestions in this direction, and we end the final chapter of the text (Chapter 9) with a suggestion toward *moral ecologies* that support moral action.

I.6.2 Overview of the Chapters

As we have presented it here, our framework to *Taking Moral Action* brings together three major themes, *contexts*, *influences*, and *processes*, which resemble the main sections of the book. Each comes with three chapters:

- *Contexts:* Chapters 1 (evolution), 2 (neuroscience), and 3 (moral ecology) each give different kinds of contexts that constrain what is possible, relevant, and appropriate in moral action. They provide broad constraints and affordances for moral action.
- *Influences:* We treat personality (Chapter 4), moral identity and the self (Chapter 5), and skills and knowledge (Chapter 6) as things that work within the contexts as

³⁰Thus, we know something about the relative position of the two groups on whatever scale is being used to measure moral identity. But only the persistent reader will find out something about the absolute position of each group on the scale. Often the “low” group is on average hovering around the midpoint of the scale, lower than the “high” group but not actually low.

³¹It is important to remember here that “people in general” is a strong claim, and one that is rarely empirically tested.

³²See Gowans (2019) for several sophisticated versions of relativism.

influences on moral action. These have often been perceived as biases, tempting the abstract autonomous reasoner away from principled judgment. We prefer the neutral term influence since it allows change either toward or away from any particular (im)moral action.

- *Processes:* Moral reason (Chapter 7), moral emotion (Chapter 8), and moral formation (Chapter 9) are three interwoven processes through which the contexts and influences produce and structure moral action, each changing and developing over time (from milliseconds to a lifetime).

Together, these constitute the nine substantive chapters of this text. It is important to note that this division into sections and chapters is largely arbitrary. The reader will see from the many cross-chapter references that they cannot easily be separated from each other. For instance, the opposition of cognition and emotion has been a guiding principle in organizing moral psychology for decades (if not millennia). We think the opposition is at least unhelpful, if not plain wrong.³³ But chapters must begin and end somewhere. We thus have separate chapters for moral reason (Chapter 7) and moral emotion (Chapter 8). Models about moral intuition are covered in Chapter 7, though they also play a central role in Chapters 6 and 8.

After reviewing, and sometimes reorganizing, the empirical literature in the area, each chapter ends with a review of the main conclusions we draw, a discussion of an historical narrative from this chapter, and a list of outstanding questions, both theoretical and empirical.

I.6.3 Chapter 1: Evolution

We first review some basic concepts and terminology in evolution. This gives us the vocabulary to describe how difficult it is to build a strong case that some particular psychological characteristic of moral action has been evolutionarily selected. We then discuss a set of six basic building blocks out of which our complex moral nature might be constructed, and review the evidence for them. We go on to suggest eleven ways that human morality is distinct from the proto-morality of other species, and ask how one might build an evolutionary argument for each aspect. We conclude that there is strong evidence that much of human morality has evolved and that there is a long (but hopeful) way to go to build evidence for how a uniquely human morality might have evolved. One implication of this grounding of morality in human evolution in small groups is that the development and expression of both morality and immorality are closely interlinked.

I.6.4 Chapter 2: Neuroscience

Instead of taking the usual approach of an anatomy-based review, we take the functional categories of morality from the chapter structure of this text and then examine how the neuroscience underlies, supports, or embodies those influences and processes in each chapter. From this perspective, the massive interconnectivity in structures in the

³³One might helpfully use the distinction to remind us of the need for a balance of influences in our planning for moral action.

brain and new methodologies that allow us to track psychological processes as they unfold over time lead us to the primary conclusion of this chapter: the reason–emotion distinction that has driven so much psychological, philosophical, and theological thinking about morality is likely false. Thus, emotion is not a bias for the individual trying to construct a moral life, it is a resource. And the dense interconnection of reason and emotion is equally a resource.

I.6.5 Chapter 3: Moral Ecology

The psychology of morality has for decades been focused on how situational influences overwhelm inclinations to moral action. This is far too narrow a view on the social embeddedness of moral action. We broaden the focus to *moral ecology*, the complex network of our social and environmental surround. To provide examples of this surround, we review intra- and inter-cultural influences, organizational and group influences, and close friendship relations, paying attention to how individuals actively shape their moral ecology. We conclude that moral ecology can be approached as a design problem (Epley & Tannenbaum, 2017) and that the psychological principles that can produce a moral community can also be used to produce cultures of corruption and the social machinery of evil.

I.6.6 Chapter 4: Personality

We present the moral personality as *adaptive* to its surroundings, *pluralistic* in its possible forms, and *integrative* of personal and situational influences. Each of the aspects are part of the answer to the question, “What does a person like me do in a situation like this?” We then review the empirical literature on personality and moral action using a three-level organization proposed by McAdams and Pals (2006): dispositional traits (e.g. conscientiousness), characteristic adaptations (e.g. authoritarianism), and self-defining life narratives (unique for each individual but with shared structure). We conclude with the somewhat paradoxical claim: It is not our virtues that make us virtuous. Rather it is the complex workings of personality that lead us to act in ways we call virtuous, and that the form this virtue takes is unique to each individual and their narrative.

I.6.7 Chapter 5: Moral Identity and the Self

In moral psychology, moral identity has been viewed as a crucial answer to the long-standing puzzle of the judgment–action gap. We review literature on moral identity that helps us to understand when and how that gap appears. We then look at research on the self, and conclude that there is no single, unitary self, and no single way that morality might be integrated into it. This multidimensionality of the self, and of self-evaluation, complicates how individuals evaluate their own consistency in bridging the gap between moral judgment and moral action. The result is an explosion of complexity in self-reflection, the conscious evaluation of the self in order to maintain consistency or integrity. We conclude that, given the human condition, perfection in self-reflection is not necessary to support an acceptable moral integrity. Instead, a modest drive to consistency in self-reflection along with modestly objective evaluations of moral action are together likely sufficient.

I.6.8 Chapter 6: Skills and Knowledge

Here, we first look at particular skills that are associated with moral action (e.g. self-reflection, self-regulation) and then review two research areas with implications for moral skill: expertise and habit. After surveying the range of human skilled behavior that expertise and habit cover, we conclude that principled, conscious, moral decision-making is probably not the primary path to moral action. Rather, the role of moral decision-making is more likely to be found in the interface between automatic action and the shift to conscious reflection when it is required.

I.6.9 Chapter 7: Moral Reason

We first ask about the nature of reason, and how various theories of reason constrain the things we can ask of moral reason. We then explore in depth the most widely recognized normative model of moral cognition, that of Lawrence Kohlberg. While looking at complaints about Kohlberg's model, we consider the problem of relativism in morality and its invitation to a cynically destructive nihilism. Here we conclude that rather than looking for independently verifiable foundations for moral certainty, the best we can hope for is a dedication to critique of our own moral actions, that of others around us, and the structure of our various moral ecologies.

We go on to look at recent research that opens up a world of implicit, nonconscious cognition. Here we find multiple parallel cognitive and emotional processes associated with moral action. Finally, we question the role of reason in supporting moral action by recognizing its dark side. We conclude that current models require an understanding of moral evaluation³⁴ that includes a careful balance among social influence, emotional evaluations, and rational reflection. We call this balancing social-emotional reflective equilibrium. This balancing is central to the process of critique implicit in Kohlberg's approach.

I.6.10 Chapter 8: Moral Emotion

We begin, as in Chapter 7, with an overview of what we might mean by emotion. We then consider the role of emotions in morality and what have been called the moral emotions. Next we discuss whether emotions are something beyond our control or if we can change or train our emotional responses. We conclude by arguing that *we could not be moral without being emotional* and that it is time to think about how one would educate the moral emotions.

I.6.11 Chapter 9: Moral Formation

Among the processes, we postulate moral formation is the one that occurs over the longest time span, indeed over a lifetime – one where the individual must integrate the influences, contexts, and processes into the life they are given. Current psychological analysis tends to fragment the processes crucial to the forming of our moral identity, and to limit analysis of them to childhood. Thus, we begin with three traditional

³⁴A more inclusive term than reason.

approaches to moral formation to highlight the unity, or at least interconnectedness, of the processes (the German *Bildung*, the Socratic-existential concept of appropriation, and monastic/spiritual formation across several religions). We draw from these to help structure the chapter and also to critique the literature and theory we cover. We then address those processes we call *evaluative* moral formation: psychological processes of how we evaluate our moral performance in the service of formation. Following that, we discuss the processes of choosing and adopting normative values to guide and motivate our action, what we call *normative* moral formation. Next, we look at the roles of education and *moral ecology* as they shape moral formation. We conclude by constructing a three-dimensional landscape for moral (trans)formation as a tool to discover variation in moral action. The dimensions are self-transcendence, openness to change, and inclusion of others (abbreviated as TOI).

I.6.12 Coda

In this short summary we draw together several of the insights from the preceding chapters:

- The centrality of critique to maintain balance in ourselves and our moral ecologies.
- A balance model of critique as social-emotional reflective equilibrium (SERE).
- The distinction between evaluative and normative moral formation and its embeddedness in moral ecology.
- The three-dimensional TOI space in which to explore the variety of moral action.

We think that any adequate theory of moral action will need to account for both the structure and the variation we present in this text and grapple with the complexity and depth of the cases and traditions included.

Finally, this book is not a disinterested essay. It is both a survey of the moral landscape and an attempt to plumb its heart (Huff & Furchert, 2023). It is committed to and written in the service of the idea that a scientific understanding of moral action can be practically beneficial, or in Lewin's (1946) terms, that it will produce *more* than just another book.

I.7 Further Reading

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019) "The Psychology of Morality: A review and analysis of empirical studies published from 1940 through 2017." This is a theoretical review of the research themes addressed by the empirical psychology of morality over the past seventy years. In addition to documenting these research themes, they conclude that the thematic areas they

describe “yield insights that are unbalanced, neglect some key features of human morality specified in influential theoretical perspectives, and are not well integrated.”

- Keane (2015) *Ethical Life: Its Natural and Social Histories*. This text provides a review of the extensive anthropological literature in morality/moral psychology. Keane covers a wide variety of cultures and venues for what he calls “ethical life” and focusses in one chapter on modern social change movements.
- Miller (2021) *Moral Psychology*. A careful and clear introduction to philosophical moral psychology that focusses on central debates: egoism, character, free will, and moral motivation.
- Shweder and Menon (2014). “Old questions for the new anthropology of morality: a commentary.” These anthropologists review the evidence for an “objective moral charter” in all cultures, and suggest this limits any claim that morality is subjective. It is a thoughtful, informed, and complex unpacking of objective or subjective nature of morality.
- Sinnott-Armstrong (2016), “The disunity of morality.” In this chapter, the philosopher Walter Sinnott-Armstrong surveys all of the approaches that claim to provide a unitary, synthesizing, or foundational view of the field and concludes that all are wanting at three levels: neurology, moral content, and moral function.

A Selective Table of Definitions of Moral Psychology

	Definition	Purpose/Context	Reference
1	Value judgments that are based on reasons, associated with the self or others as good or bad, that are universalizable, subject to consistency pressure, and considered to be objective in nature	To define “a set of criteria which correspond to enlightened ordinary language usage of the concept of a moral judgment” that allows “psychological study of the moral in both a descriptive and normative sense” (p. 6)	Summary of Kohlberg (1958) pp. 8–12
2	“Prescriptive (ought) judgments, found in all cultures, based in conscious rational application of general rules and principles with the central function of resolving social conflicts over claims or rights”	A synopsis and defense of the full theory and research program. It retains the original goals listed in definition 1 and grounds its objectivity in conscious rational judgment	Levine, Kohlberg, & Hower (1985) pp. 97–98

	Definition	Purpose/Context	Reference
3	“theories of moral personhood cannot avoid endorsing goals for <i>how individuals should reason</i> right from wrong and <i>what individuals should value</i> as being good” ³⁵	In a section titled “defining moral behavior” they propose that important conceptual work is exploring “which non-neutral conception of the good is most appropriate” p. 338	Frimer & Walker (2008) p. 337
4	“a sense of morality is ... a mental phenomenon that consists in thoughts and feelings about rights and duties, good and bad character traits (virtues and vices), and right and wrong motives and behaviors” It “contains feelings and standards of conduct, may be embedded in religion, and some aspects are universal across cultures”	A definition that allows us to “identify the mental mechanisms that produce [a sense of morality] and explain how they were selected and refined in the course of human evolution” (p. 149)	Krebs (2008) p. 150
5	“evaluations (good versus bad) of the actions or character of a person that are made with respect to a set of virtues held by a culture or subculture to be obligatory”	To construct a model of moral evaluations (vs. rational judgments) that includes intuition and “emphasizes instead the importance of social and cultural influences” p. 814 ³⁶	Haidt (2001) p. 817
6	“moral emotions are those emotions that are linked to the interests or welfare either of society as a whole or at least of persons other than the judge or agent”	An early attempt to systematize the burgeoning literature in the moral emotions	Haidt (2003) p. 853
7	“Moral systems are interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make social life possible”	Construction of a questionnaire to measure aspects of Moral Foundations Theory	Graham et al. (2011) p. 368

³⁵See Chapter 9, where Frimer and Walker’s research program is described.

³⁶Note that the purpose of this definition from Haidt (2001) is an explicit rejection of Kohlberg’s goal of adopting rationalism in order to reject relativism. See Chapter 7 for the importance of rationalism in Kohlberg’s system as a defense against relativism.

	Definition	Purpose/Context	Reference
8	Moral mandates are experienced as “Absolute/ Universal, objective, authority independent, motivating, self-justifying, strong emotions”	A theoretical attempt to describe a domain of attitudes, <i>moral mandates</i> , that “differ in form and implication from otherwise strong but nonmoral attitudes” p. 278	Skitka (2010) p. 278
9	“prescriptive judgments of justice, rights, and welfare pertaining to how people ought to relate to each other”	To distinguish “two social domains: convention and morality” p. 3	Turiel (1983) p. 3
10	Ethical life refers to those aspects of people’s everyday actions “as well of their sense of themselves and of other people (and sometimes entities such as gods or animals) that are oriented with reference to values and ends that are not in turn defined as the means to some further ends”	An anthropological “natural and social history” of ethical life to explain the ways that ordinary life is shot through with moral evaluation in all cultures	Keane (2015) p. 4
11	“Virtuous individuals, by this account, would be those for whom moral categories are chronically accessible for appraising and interpreting the social landscape”	An attempt to explain “traits, virtues and other dispositional features of moral character” that take advantage of work in social cognition and expertise and that can inform philosophical ethics (p. 30)	Lapsley & Narvaez (2005) p. 30
12	“Moral behavior involves a state of being” toward oneself and the world. It involves being “fully alive to the situation, to life in one’s surrounds (human and nonhuman), to connectedness, interdependence, and influence with that life”	Triune Ethics Meta-Theory proposes an ideal state of moral functioning based on our evolved capacities for care and openness ³⁷	Narvaez (2016) p. 34 - 36
13	Morality is a “specific form of functioning of each person as a whole”	A review of the relationship of moral psychology and personality psychology	Cervone & Tripathi (2009) p. 31, citing Blasi (2004) p. 336

³⁷See Chapter 6 for an idea of what capacities this “ideal state” entails.

Definition	Purpose/Context	Reference
14 “I will count a judgment as a moral judgment if and only if [millions of] people who make that judgment do, or would think of it as similar enough in important respects to judgments that they take to be exemplars or paradigms of moral judgments”	From a philosophical paper whose thesis is that there is no foundational unity to morality at three levels: the neural level, the level of moral content, or the level of moral function	Sinnott-Armstrong (2016) p. 332
15 We define moral reasoning as transitions in thought in accordance with endorsed moral principles ... In turn, we define moral principles as prescriptive principles concerning others’ welfare, rights, equality, fairness, and justice	From a review paper that argues that, from childhood to adulthood, moral reasoning is associated with taking moral action in both individual and cultural situations	Killen & Dahl (2021, pp. 1210-1211)
16 Morality-as-cooperation [... includes] helping kin, helping your group, reciprocating, being brave, deferring to superiors, dividing disputed resources, and respecting prior possession	From an anthropological paper looking for cooperation as a moral cultural universal across 60 societies	Curry, Mullins, & Whitehouse (2019)

References

Adelman, C. (1993). Kurt Lewin and the origins of action research. *Educational Action Research*, 1(1), 7–24. doi:10.1080/0965079930010102

Allport, G. W. (1954/1988). *The Nature of Prejudice*. Reading, MA: Addison-Wesley.

Atkins, R., Hart, D., & Donnelly, T. M. (2005). The association of childhood personality type with volunteering during adolescence. *Merrill-Palmer Quarterly*, 51(2), 145–162. doi:10.1353/mpq.2005.0008

Augustine, M. J. (1986). Zen and Benedictine monks as mythopoeic models of nonegocentered worldviews and lifestyles. *Buddhist-Christian Studies*, 6, 23–49.

Barrett, L. F., Adolphs, R., Marsella, S. et al. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68.

Bebeau, M. J. (1994). Influencing the moral dimensions of dental practice. In J. R. Rest & D. Narvaez (Eds.), *Moral Development in the Professions: Psychology and Applied Ethics* (pp. 121–146). Hillsdale, NJ: Lawrence Erlbaum.

Bebeau, M. J., & Monson, V. E. (2012). Professional Identity Formation and Transformation across the Life Span. 135–162. doi:10.1007/978-94-007-1724-4_7

- Bebeau, M. J., & Thoma, S. J. (2012). Moral motivation in different professions. In K. Heinrichs, F. Oser, & T. Lovat (Eds.), *Handbook of Moral Motivation* (pp. 436–460). Sense Publishers.
- Beitsch, R. (July 20, 2016). Victims and offenders come together to make their own justice. Retrieved from <http://www.pewtrusts.org/en/research-and-analysis/blogs/state-line/2016/07/20/victims-and-offenders-come-together-to-make-their-own-justice>
- Bellah, R. N. (2011). *Religion in Human Evolution: From the Paleolithic to the Axial Age*. Cambridge, MA: Belknap.
- Berry, Z., Lewis, N. A., & Sowden, W. J. (2021). The Double-Edged Sword of Loyalty. *Current Directions in Psychological Science*, 30(4), 321–326. doi:10.1177/09637214211010759
- Beuys, B. (2012). *Sophie Scholl Biographie*. Munich: Carl Hanser Verlag.
- Bidgood, J. (2021, July 5). Aaron van Langevelde's speech about the 2020 election: "We were asked to take power we didn't have." *Boston Globe*.
- Blasi, A. (2004). Moral functioning: Moral understanding and personality. In D. Lapsley & D. Narvaez (Eds.), *Moral Development, Self, and Identity* (pp. 189–212). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bloom, P. (2012). Religion, morality, evolution. *Annual Review of Psychology*, 63, 179–199. doi:10.1146/annurev-psych-120710-100334
- Borenstein, N. S. (1989). My life as a NATO collaborator. *Bulletin of the Atomic Scientists*, April, 14–20.
- Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspectives on Psychological Science*, 15(4), 978–1010. doi:10.1177/1745691620917336
- Braithwaite, J. (1989). Criminological theory and organizational crime. *Justice Quarterly*, 6(3), 333–358.
- Braithwaite, J. (2002). *Restorative Justice & Responsive Regulation*. New York: Oxford University Press.
- Braithwaite, J. (2014). Evidence for restorative justice. *Vermont Bar Journal*, Summer, 18–22.
- Braithwaite, J. (2016). Braithwaite (2016) Learning to scale up restorative justice. In K. Clamp (Ed.), *Restorative Justice in Transitional Settings* (pp. 173–189). Abingdon: Routledge.
- Bucar, E. M. (2018). Islamic virtue ethics. In N. E. Snow (Ed.), *Oxford Handbook of Virtue*. Oxford: Oxford University Press.
- Callaway, K., & Strawn, B. (2020). Experimental theology: Theological anthropology and the psychological sciences. *Journal of Psychology and Theology*, 48, 3–17. doi:10.1177/0091647119854117
- Cannon, S. F. (1978). *Science in Culture*. New York: Dawson and Science History Publications.
- Cervone, D. (2021). Five paths to personality coherence: Integrative implications of the Knowledge-and-Appraisal Personality Architecture. *European Journal of Personality*, 36(3), 319–346. doi:10.1177/08902070211015599
- Cervone, D., & Tripathi, R. (2009). The moral functioning of the person as a whole: On moral psychology and personality science. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 30–51). New York: Cambridge University Press.
- Ciecuch, J., Schwartz, S. H., & Davidov, E. (2015). The social psychology of values. In J. D. Wright (Ed.) *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., Vol. 25, pp. 41–46) Oxford: Elsevier. doi:10.1016/b978-0-08-097086-8.25098-8
- Clark, C. J., Liu, B. S., Winegard, B. M. et al. (2019). Tribalism is human nature. *Current Directions in Psychological Science*, 28(6), 587–592. doi:10.1177/0963721419862289
- Colby, A., & Damon, W. (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. New York: Free Press.

- Colby, A., & Damon, W. (2015). *The Power of Ideals: The Real Story of Moral Choice*. Oxford: Oxford University Press.
- Collins, H., & Pinch, T. (2006). *Dr. Golem: How to Think About Medicine*. Chicago, IL: University of Chicago Press.
- Comer, D. R., & Sekerka, L. E. (2018). Keep calm and carry on (ethically): Durable moral courage in the workplace. *Human Resource Management Review*, 28(2), 116–130. doi:10.1016/j.hrmr.2017.05.011
- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1), 47–69. doi:10.1086/701478
- Damon, W., & Colby, A. (2013). Why a true account of human development requires exemplar research. *New Directions in Child and Adolescent Development*, 2013(142), 13–25. doi:10.1002/cad.20046
- Darley, J. M. (2001). Citizen's sense of justice and the legal system. *Current Directions in Psychological Science*, 10(1), 10–13.
- Darley, J. M., & Batson, C. D. (1973). "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27(1), 100–108.
- Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, 24(6), 659–683.
- Dettelbach, M. (2001). Alexander von Humboldt between enlightenment and romanticism. *Northeastern Naturalist*, 8, 9–20.
- DeWall, C. N., Giancola, P. R., McCullough, M. E. et al. (2010). *Does Religion Inhibit or Increase Impulsive Violence? Evidence that Religiosity Buffers People from Alcohol-Related Violence*. University of Kentucky.
- DeYoung, C. G., & Gray, J. R. (2009). Personality Neuroscience: Explaining individual differences in affect, behavior, and cognition. In P. J. Corr & G. Matthews (Eds.), *The Cambridge Handbook of Personality Psychology* (pp. 323–346). Cambridge: Cambridge University Press.
- DeYoung, C. G., Hirsh, J. B., Shane, M. S. et al. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science*, 1–9.
- Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Doris, J. M., & Stich, S. P. (2005). As a matter of fact: Empirical perspectives on ethics. In F. Jackson & M. Smith (Eds.) *The Oxford Handbook of Contemporary Philosophy* (pp. 114–152) Oxford: Oxford University Press.
- Dovidio, J. F., Piliavin, J. A., Gaertner, S. L. et al. (1991). The arousal: Cost–reward model and the process of intervention: A review of the evidence. In Clark, Margaret S. *Prosocial Behavior. Review of Personality and Social Psychology* (vol. 1912, pp. 1986–1118). Thousand Oaks, CA: Sage Publications.
- Dovidio, J. F., Piliavin, J. A., Schroeder, D. A. et al. (2006). *The Social Psychology of Prosocial Behavior*. New York: Taylor & Francis.
- Ellemers, N., van der Toorn, J., Paunov, Y. et al. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*, 23(4), 332–366. doi:org/10.1177/1088868318811
- Epley, N., & Tannenbaum, D. (2017). Treating ethics as a design problem. *Behavioral Science & Policy*, 3(2), 73–84.
- Everett, J. A. C., Haque, O. S., & Rand, D. G. (2016). How good is the samaritan, and why? An experimental investigation of the extent and nature of religious prosociality using economic games. *Social Psychological and Personality Science*, 7(3), 248–255. doi:10.1177/1948550616632577
- Fleischmann, A., Lammers, J., Conway, P. et al. (2017). Paradoxical effects of power on moral thinking: Why power both increases and decreases deontological and utilitarian moral

- decisions. *Social Psychological and Personality Science*, 10(1), 110–120. doi:10.1177/1948550617744022
- Fowers, B. J., Carroll, J. S., Leonhardt, N. D. et al. (2020). The emerging science of virtue. *Perspectives on Psychological Science*, 16(1), 118–147. doi:10.1177/1745691620924473
- Frankl, V. (1959). *Man's Search for Meaning*. New York: Washington Square Press.
- Frankl, V. (1997). *Man's Search for Ultimate Meaning*. New York: Insight Books.
- Frankl, V. (2018). *Der leidende Mensch: Anthropologische Grundlagen der Psychotherapie*. Göttingen: Hogrefe.
- Frimer, J. A., Gaucher, D., & Schaefer, N. K. (2014). Political conservatives affinity for obedience to authority is loyal, not blind. *Personality and Social Psychology Bulletin*, 40(9), 1205–1214.
- Frimer, J. A., & Walker, L. J. (2008). Towards a new paradigm of moral personhood. *Journal of Moral Education. Special Issue: Towards an integrated model of moral functioning*, 37(3), 333–356. doi:10.1080/03057240802227494
- Fung, K. (2022, August 17). Politics: Read Liz Cheney's concession speech as congresswoman ousted in Wyoming. *Newsweek*.
- Furchert, A. (2012). *Das Leiden fassen: zur Leidensdialektik Søren Kierkegaards*. Freiburg: Karl Alber.
- Furchert, A. (2022). Bringing God into Existence. Kierkegaard's philosophy of appropriation revisited. Retrieved from https://www.academia.edu/55140937/Bringing_God_into_Existence_Kierkegaards_philosophy_of_appropriation_revisited
- Gadamer, H.-G. (1975). *Wahrheit und Methode. Grundzüge einer philosophischen Hermeneutik*. Tübingen: J. C. B. Mohr.
- Gadamer, H.-G. (1996). *The Enigma of Health: The Art of Healing in a Scientific Age*. Stanford, CA: Stanford University Press.
- Gamez-Djokic, M., & Molden, D. (2016). Beyond affective influences on deontological moral judgment: the role of motivations for prevention in the moral condemnation of harm. *Personality & Social Psychology Bulletin*, 42(11), 1522–1537. doi:10.1177/0146167216665094
- Gardner, A., & Firozi, P. (2021, January 5). Here's the full transcript and audio call between Trump and Raffensperger. *Washington Post*.
- Gibbs, J. C., Basinger, K. S., Grime, R. L. et al. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review*, 27(4), 443–500. doi:10.1016/j.dr.2007.04.001
- Gill, M. J., & Cerce, S. C. (2017). He never willed to have the will he has: Historicist narratives, “civilized” blame, and the need to distinguish two notions of free will. *Journal of Personality and Social Psychology*, 112(3), 361–382. doi:10.1037/pspa0000073
- Gill, M. J., & Getty, P. D. (2016). On shifting the blame to humanity: Historicist narratives regarding transgressors evoke compassion for the transgressor but disdain for humanity. *British Journal of Social Psychology*, 55(4), 773–791. doi:10.1111/bjso.12159
- Ginges, J., Hansen, I., & Norenzayan, A. (2009). Religion and support for suicide attacks. *Psychological Science*, 20(2), 224–230. doi:10.1111/j.1467-9280.2009.02270.x
- Gowans, C. (2019). Moral relativism. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. Retrieved from <https://plato.stanford.edu/archives/sum2019/entries/moral-relativism/>
- Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review*, 14(1), 140–150.
- Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. R. Shaver & M. Mikulincer (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil* (pp. 11–31). New York: American Psychological Association Books.

- Graham, J., Nosek, B. A., Haidt, J. et al. (2011). Mapping the Moral Domain. *Journal of Personality and Social Psychology*, 101, 366–385.
- Greene, J. D. (2014). The cognitive neuroscience of moral judgment and decision making. In M. S. Gazzaniga & G. R. Mangun (Eds.), *The Cognitive Neurosciences* (5th ed., pp. 1013–1024). Cambridge, MA: MIT Press.
- Greene, J. D. (2015). The rise of moral cognition. *Cognition*, 135, 39–42. doi:10.1016/j.cognition.2014.11.018
- Grube, J. A., & Piliavin, J. A. (2000). Role identity, organizational experiences, and volunteer performance. *Personality and Social Psychology Bulletin*, 26(9), 1108–1119.
- Habashi, M. M., Graziano, W. G., & Hoover, A. E. (2016). Searching for the prosocial personality: A big-five approach to linking personality and prosocial behavior. *Personality and Social Psychology Bulletin*, 42(9), 1177–1192. doi:10.1177/0146167216652859
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, Scherer, K. R., & Goldsmith, H. H. (Ed.), *Handbook of Affective Sciences* (pp. 852–870). Oxford: Oxford University Press.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998.
- Haidt, J., & Kesebir, S. (2010). Morality. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (5th ed., pp. 787–832). Hoboken, NJ: Wiley.
- Hamilton, N. W., & Monson, V. E. (2012). Ethical professionalism (trans) formation: themes from interviews about professionalism with exemplary lawyers. *Santa Clara Law Review*, 52(3), 921–970.
- Harrington, A. (2000). In defence of Verstehen and Erklären: Wilhelm Dilthey's ideas concerning a descriptive and analytical psychology. *Theory & Psychology*, 10(4), 435–451.
- Harris, C. E. (2008). The good engineer: Giving virtue its due in engineering ethics. *Science and Engineering Ethics*, 14(2), 153–164.
- Harris, N. (2006). Reintegrative shaming, shame, and criminal justice. *Journal of Social Issues*, 62(2), 327–346.
- Hart, D. (2005). Adding identity to the moral domain. *Human Development*, 48, 257–261.
- Hart, D., Atkins, R., & Southerland, N. (2006). Community service and moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 633–656). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hart, D., & Matsuba, M. K. (2007). The development of pride and moral life. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The Self-Conscious Emotions: Theory and Research* (pp. 114–133). New York: Guilford Press.
- Hartnagel, T. (2005). *Sophie Scholl, Fritz Hartnagel; Damit wir uns nicht verlieren: Briefwechsel 1937–1943*. Frankfurt: S. Fischer.
- Hartshorne, H., & May, M. A. (1928). *Studies in the Nature of Character: Studies in Deceit* (vol. 1). New York: Macmillan.
- Hartshorne, H., May, M. A., & Maller, J. B. (1929). *Studies in the Nature of Character. II. Studies in Service and Self-Control*. (Vol. 2). New York: Macmillan.
- Hartshorne, H., May, M. A., & Shuttlesworth, F. K. (1930). *Studies in the Nature of Character. III. Studies in the Organization of Character* (vol. 3). New York: Macmillan.
- Heiphetz, L., & Cushman, F. (2021). Introduction to morality as a hub: Connections within and beyond social cognition. *Social Cognition*, 39(1), 1–3.
- Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Hoffman, L., & Ellis, M. S. (2019). Ultimate concern. In D. Leeming (Ed.), *Encyclopedia of Psychology and Religion* (pp. 1–4). New York: Springer.

- Huff, C., & Barnard, L. K. (2009). Good computing: moral exemplars in the computing profession. *IEEE Technology and Society Magazine*, (Fall), 47–54.
- Huff, C., Barnard, L. K., & Frey, W. (2008a). Good computing: A pedagogically focused model of virtue in the practice of computing (part 1). *Journal of Information, Communication and Ethics in Society*, 6(3), 246–278.
- Huff, C., Barnard, L. K., & Frey, W. (2008b). Good Computing: a pedagogically focused model of virtue in the practice of computing (part 2). *Journal of Information, Communication & Ethics in Society*, 6, 284–316.
- Huff, C., & Furchert, A. (2014). Toward a pedagogy of ethical practice: the application problem. *Communications of the ACM*, 27(7), 25–27.
- Huff, C., & Furchert, A. (2016). Embracing the complexity of moral action. In D. Bertini (Ed.), *Moral Heteronomy. History, Proposals, Arguments, a special issue of Dialegestha*.
- Huff, C., & Furchert, A. (2023). Existential appropriation in a digital age: The heart at the heart of digital ethics. In S. Rogerson (Ed.), *Handbook of Digital Social Science*. Cheltenham: Edgar Elgar.
- Huff, C., & Gaasedelen, O. (2011). Moral psychology. *Oxford Bibliographies Online*. Retrieved from <http://www.oxfordbibliographiesonline.com/>
- Huff, C., & Martin, C. D. (1995). Computing consequences: A framework for teaching ethical computing. *Communications of the ACM*, 38(12), 75–84.
- Hulsey, T. L., & Hampson, P. J. (2014). Moral expertise. *New Ideas in Psychology*, 34(0), 1–11. doi:10.1016/j.newideapsych.2014.02.001
- Hursthouse, R., & Pettigrove, G. (2016). Virtue ethics. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University.
- Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: two faces of moral regulation. *Journal of Personality and Social Psychology*, 96(3), 521–537. doi:10.1037/a0013779
- Jaspers, K. (1938). *Existenzphilosophie*. Berlin: de Gruyter.
- Jaspers, K. (1958). Der Arzt im technischen zeitalter. *Klinische Wochenschrift*, 36(22), 1037–1043.
- Jaspers, K. (1973). *Allgemeine Psychopathologie*. New York: Springer.
- Kant, I. (1785/2011). *Groundwork of the Metaphysics of Morals: A German–English Edition* (M. Gregor & J. Timmermann, Trans.). Cambridge: Cambridge University Press.
- Karl, J. (2022 August 21). TRANSCRIPT: ABC News’ Jonathon Karl interviews Rep. Liz Cheney. ABC News.
- Keane, W. (2015). *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton University Press.
- Keller, E. F. (1983). *A Feeling for the Organism: The Life and Work of Barbara McClintock*. New York: W. H. Freeman and Co.
- Kierkegaard, S. (1849/2004). *The Sickness Unto Death* (A. Hannay, Trans.). New York: Penguin.
- Kilcourse, G. (1996). Thomas Merton and racism: “Letters to a White Liberal” reconsidered. Paper presented at the First general meeting of the Thomas Merton Society of Great Britain and Ireland. Southampton, England. Retrieved from <http://www.thomasmerton-society.org/kilcour.htm>
- Killen, M., & Dahl, A. (2021). Moral reasoning enables developmental and social change. *Perspectives on Psychological Science*, 16(6), 1209–1225. doi:10.1177/1745691620964076
- Kohlberg, L. (1958). The development of modes of moral thinking and choice in the years 10 to 16. (PhD Dissertation). Chicago, IL: University of Chicago Press.
- Koltko-Rivera, M. E. (2006). Rediscovering the later version of Maslow’s hierarchy of needs: Self-transcendence and opportunities for theory, research, and unification. *Review of General Psychology*, 10(4), 302–317. doi:10.1037/1089-2680.10.4.302

- Koole, S. L., McCullough, M. E., Kuhl, J. et al. (2010). Why religion's burdens are light: from religiosity to implicit self-regulation. *Personality and Social Psychology Review*, 14(1), 95–107. doi:10.1177/1088868309351109
- Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science*, 3(3), 149.
- Kuhl, J., Quirin, M., & Koole, S. L. (2021). The functional architecture of human motivation: Personality systems interactions theory. In A. Elliot (Ed.), *Advances in Motivation Science* (vol. 8, pp. 1–62). New York: Elsevier.
- Lapsley, D. K. (2008). Moral Self-Identity as the aim of education. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of Moral and Character Education* (pp. 30–52). New York: Routledge.
- Lapsley, D. K., & Lasky, B. (2001). Prototypic moral character. *Identity: An International Journal of Theory and Research*, 1(4), 345–363.
- Lapsley, D. K., & Narvaez, D. (2005). Moral psychology at the crossroads. In D. K. Lapsley & F. C. Power (Eds.), *Character Psychology and Character Education* (pp. 18–35). South Bend, IN: University of Notre Dame Press.
- Latane, B., & Darley, J. M. (1970). *The Unresponsive Bystander: Why Doesn't He Help?* Englewood Cliffs, NJ: Prentice Hall.
- Levine, C., Kohlberg, L., & Hower, A. (1985). The current formulation of Kohlberg's theory and a response to critics. *Human Development*, 28, 94–100.
- Lewin, K. (1945). The Research Center for Group Dynamics at Massachusetts Institute of Technology. *Sociometry*, 8(2), 126–136.
- Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, 2(4), 34–46.
- Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created "social climates." *Journal of Social Psychology*, 10, 271–299.
- Luhrmann, T., & Weisman, K. (2022). Porosity is the heart of religion. *Current Directions in Psychological Science*. doi:10.1177/09637214221075285
- Lundin, R. (2004). *Emily Dickinson and the Art of Belief*. Grand Rapids, MI: Eerdmans.
- MacKenzie, M. (2017). Buddhism and the virtues. In N. E. Snow (Ed.), *The Oxford Handbook of Virtue*. Oxford: Oxford University.
- Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology*, 72, 293–318. doi:10.1146/annurev-psych-072220-104358
- Malpas, J. (2018). Hans Georg Gadamer. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University.
- Maslow, A. H. (1969). The farther reaches of human nature. *The Journal of Transpersonal Psychology*, 1(1), 1–9.
- McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, 61(3), 204–217.
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509–516.
- Merton, T. (1948/1998). *The Seven Story Mountain*. New York: Houghton Mifflin Harcourt.
- Merton, T. (1964). *Seeds of Destruction*. New York: Farrar, Straus and Giroux.
- Merton, T. (1966). *Conjectures of a Guilty Bystander*. New York: Crown.
- Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371–378.
- Milgram, S. (1974). *Obedience to Authority: An Experimental View*. New York: Harper and Row.
- Miller, C. B. (2021). *Moral Psychology*. New York: Cambridge University Press.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.
- Mischel, W., & Morf, C., C. (2012). The self as a psycho-social dynamic processing system. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of Self and Identity* (2nd ed., pp. 21–49). New York: Guilford Press.

- Narvaez, D. (2005). The neo-Kohlbergian tradition and beyond: Schemas, expertise, and character. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (vol. 51, pp. 119–163). Lincoln, NE: University of Nebraska Press.
- Narvaez, D. (2016). *Embodied Morality: Protectionism, Engagement and Imagination*. New York: Palgrave Pivot.
- Narvaez, D., & Rest, J. R. (1995). The four components of acting morally. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Moral Development: An Introduction* (pp. 385–399). New York: Allyn & Bacon.
- Oliner, S. P., & Oliner, P. M. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press
- Penner, L. A., Dovidio, J. F., Piliavin, J. A. et al. (2004). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, 5, 365–392.
- Penner, L. A., & Orom, H. (2010). Enduring goodness: A person-by-situation perspective on prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature* (pp. 55–72). Washington, DC: American Psychological Association.
- Plaisance, P. (2014a). *Virtue in Media: The Moral Psychology of Excellence in News and Public Relations*. London: Routledge.
- Plaisance, P. (2014b). Virtue in media: The moral psychology of U.S. exemplars in news and public relations. *Journalism & Mass Communication Quarterly*, 91(2), 308–325. doi:10.1177/1077699014527460
- Pritchard, M. S. (1998). Professional responsibility: Focusing on the exemplary. *Science and Engineering Ethics*, 4(2), 215–233.
- Pritchard, M. S. (2001). Responsible engineering: the importance of character and imagination. *Science and Engineering Ethics*, 7(3), 391–402.
- Raffensperger, B. (2021). *Integrity Counts*. Brentwood, TN: Forefront Books.
- Rambachan, A. (2006). *The Advaita Worldview: God, World, and Humanity*. New York: State University of New York Press.
- Roberts, B. W., & Mroczek, D. K. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, 17(1), 31–35.
- Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and Conservatives Share the same harm-based moral template. *Personality and Social Psychology Bulletin*, 41(8), 1147–1163. doi:10.1177/0146167215591501
- Schwartz, B. (2010). The limits of cost-benefit calculation: Commentary on Bennis, Medin, & Bartels (2010). *Perspectives on Psychological Science*, 5(2), 203–205. doi:10.1177/1745691610362355
- Shanley, B. J. (2008). Aquinas's exemplar ethics. *The Thomist*, 72, 346–369.
- Shenhav, A., & Greene, J. D. (2014). Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *Journal of Neuroscience*, 34(13), 4741–4749. doi:10.1523/JNEUROSCI.3390-13.2014
- Shweder, R. A. (2012). Relativism and universalism. In D. Fassin (Ed.), *A Companion to Moral Anthropology* (pp. 85–102). New York: Wiley-Blackwell.
- Shweder, R. A., & Menon, U. (2014). Old questions for the new anthropology of morality: a commentary. *Anthropological Theory*, 14(3), 356–370.
- Sinnott-Armstrong, W. (2016). The disunity of morality. In S. M. Liao (Ed.), *Moral Brains: The Neuroscience of Morality* (pp. 331–353). New York: Oxford University Press.
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, 4(4), 267–281.
- Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, 88(6), 895–917. doi:10.1037/0022-3514.88.6.895

1 Taking Moral Action

- Snow, N. E. (2010). *Virtue as Social Intelligence: An Empirically Grounded Theory*. New York: Routledge.
- Snyder, M., & Clary, E. G. (2004). Volunteerism and the generative society. In E. de St. Aubin, D. McAdams, & T. C. Kim (Eds.), *The Generative Society: Caring for Future Generations* (pp. 221–237). Washington, DC: American Psychological Association.
- Snyder, M., & Omoto, A. M. (2008). Volunteerism: Social issues perspectives and social policy implications. *Social Issues and Policy Review*, 2(1), 1–36.
- Stich, S. P., Doris, J. M., & Roedder, E. (2010). Altruism. In J. M. Doris (Ed.), *The Moral Psychology Handbook* (pp. 147–205). New York: Oxford University Press.
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171. doi:10.1016/j.cognition.2013.12.005
- Stuewig, J., Tangney, J. P., Heigel, C. et al. (2010). Shaming, blaming, and maiming: Functional links among the moral emotions, externalization of blame, and aggression. *Journal of Research in Personality*, 44(1), 91–102. doi:10.1016/j.jrp.2009.12.005
- Stuewig, J., Tangney, J. P., Mashek, D. et al. (2009). The Moral Emotions, Alcohol Dependence, and HIV Risk Behavior in an Incarcerated Sample. *Substance Use & Misuse*, 44(4), 449–471. doi:10.1080/10826080802421274
- Stukas, A. A., Snyder, M., & Clary, E. G. (2016). Understanding and encouraging volunteerism and community involvement. *The Journal of Social Psychology*, 156(3), 243–255. doi:10.1080/00224545.2016.1153328
- Tangney, J. P., Mashek, D., & Stuewig, J. (2007). Working at the social-clinical-community-criminology interface: The George Mason University inmate study. *Journal of Social and Clinical Psychology*, 26(1), 1–21.
- Tangney, J. P., Stuewig, J., & Hafez, L. (2011). Shame, guilt, and remorse: implications for offender populations. *The Journal of Forensic Psychiatry and Psychology*, 22(5), 706–723.
- Tangney, J. P., Stuewig, J., & Martinez, A. G. (2014). Two faces of shame: The roles of shame and guilt in predicting recidivism. *Psychological Science*, 25(3), 799–805. doi:10.1177/0956797613508790
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58, 345–372.
- Taylor, S. E., Klein, L. C., Lewis, B. P. et al. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, 107(3), 411–429. doi:10.1037/0033-295x.107.3.411
- Thagard, P. (2014). Explanatory identities and conceptual change. *Science & Education*, 23(7), 1531–1548. doi:10.1007/s11191-014-9682-1
- Thalhammer, K. (2001). I'll take the high road: Two pathways to altruistic political mobilization against regime repression in Argentina. *Political Psychology*, 22(3), 493–519.
- Thalhammer, K., O'Loughlin, P. L., Glazer, M. P. et al. (2007). *Courageous Resistance: The Power of Ordinary People*. Basingstoke: Palgrave MacMillan.
- Thomas Merton Center. (n.d.). Thomas Merton's life and work. Retrieved from <http://merton.org/chrono.aspx#links>
- Tiberius, V. (2014). *Moral Psychology: A Contemporary Introduction*. New York: Routledge.
- Tillich, P. (1951). *Systematic Theology, Vol. 1, Reason and Revelation. Being and God*. Chicago, IL: University of Chicago Press.
- Tillich, P. (1952). *The Courage to Be*. New Haven: Yale.
- Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.
- Vallor, S. (2014). Experimental virtue: perceptual responsiveness and the praxis of scientific observation. In A. Fairweather (Ed.), *Virtue Epistemology Naturalized: Bridges Between Virtue Epistemology and Philosophy of Science* (pp. 269–290). New York: Springer.

- van Wyhe, J. (2002). Humboldt's personal narrative and its influence on Darwin. Retrieved from http://darwin-online.org.uk/EditorialIntroductions/Chancellor_Humboldt.html
- Vonash, A. J., Reynolds, T., Winegard, B. M. et al. (2017). Death before dishonor: Incurring costs to protect moral reputation. *Social Psychological and Personality Science*, 9(5), 604–613. doi:10.1177/1948550617720271
- Walker, C. (1995). Karl Jaspers and Edmund Husserl IV: Phenomenology as empathic understanding. *Philosophy, Psychiatry, & Psychology*, 2(3), 247–266.
- Walker, D. I., & Thoma, S. J. (2017). Moral and character education. In G. Noblit (Ed.), *Oxford Research Encyclopedia of Education*. Retrieved from <https://oxfordre.com/education/display/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-119?rskey=NKhAB8&result=1>
- Walker, L. J. (2004). Progress and prospects in the psychology of moral development. *Merrill-Palmer Quarterly*, 50(4), 546–557.
- Walker, L. J., & Frimer, J. A. (2008). Being good for goodness' sake: Transcendence in the lives of moral heroes. In F. Oser & W. Veugelers (Eds.), *Getting Involved: Global Citizenship Development and Sources of Moral Values* (pp. 309–326). Rotterdam: Sense.
- Walker, L. J., & Frimer, J. A. (2015). Developmental trajectories of agency and communion in moral motivation. *Merrill-Palmer Quarterly*, 61(3), 412–439.
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, 86(4), 629–647. doi:10.1037/0022-3514.86.4.629
- Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review*, 8(3), 281–307. doi:10.1207/s15327957pspr0803_4
- Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, 21(3), 253–277. doi:10.1177/1088868316652279
- Ysseldyk, R., Matheson, K., & Anisman, H. (2010). Religiosity as identity: Toward an understanding of religion from a social identity perspective. *Personality and Social Psychology Review*, 14(1), 60–71. doi:10.1177/1088868309349693
- Zagzebski, L. (2013). Moral exemplars in theory and practice. *Theory and Research in Education*, 11(2), 193–206. doi:10.1177/1477878513485177
- Zagzebski, L. (2017). *Exemplarist Moral Theory*. Oxford: Oxford University Press.
- Zuckerman, M. (1995). Good and bad humors: Biochemical bases of personality and its disorders. *Psychological Science*, 6(6), 325–332.

Final Proofs

Part I

Contexts

Final Proofs

Final Proofs

Evolution

1.1 The Social Brain

The human brain is among the largest relative to body size in mammals (Herculano-Houzel, 2009), but more surprising is that in the evolutionary time scale, the brain has recently (in the last million years or so) shown rapid increases in size (Adolphs, 2009). This poses a puzzle: Why such a massive increase in such a relatively short time? One widely accepted answer to this puzzle is the social brain hypothesis (Adolphs, 2009; Byrne & Whiten, 1988; Whiten & van Schaik, 2007). This is the idea that the rapid development of the brain was in response to¹ increases in the complexity of the *social problems* that were necessary for early hominids to solve. These problems include many of those we will review in this chapter: various forms of cooperation, deception, status hierarchy negotiation, altruistic punishment, and anticipatory fear of punishment. It is during this rapid spate of development that humans became a species able to enact a complex (im)morality.

Philosophers as early as Aristotle (1941) recognized the importance of the social nature of humans. His discussion of humans as *zōon politikon*² in the early chapters of the *Politics* makes it clear that our social nature (as opposed to that of bees or cranes) is intimately related to our concern for justice. For Aristotle and many thinkers that followed, the connection between our social nature and our moral nature is intimate and inextricable. We will find this to be true in our evolutionary history as well.

We know today that moral judgment and action involve a rich, complex interaction of emotion, inference, automatic and controlled processing, strategic planning, etc. A core assumption of evolutionary psychology is that the capacities for these activities evolved under the pressure of selection from the social environment and together constitute our distinct form of moral/social organization as the human species (Krebs, 2008).

¹Or at least in coevolutionary tandem with, see Section 2.4.

²Perhaps best translated as *political beings*. Aristotle's *Politics* is about the role that the community and its lawgivers should play in bringing about the virtuous life in the citizens. Thus, morality is closely linked to our social nature.

To understand how these capacities might have evolved, we first need to understand some basic concepts and terminology in evolution. That will give us the vocabulary to describe how difficult it is to build a strong case that some particular psychological characteristic has been evolutionarily selected. After we see what a good evolutionary argument looks like, we will then review a set of six basic building blocks out of which our complex moral nature might be constructed, and review the evidence for them. Finally, we will look at a set of eleven ways that human morality is distinct from the proto-morality of other species, and ask how one might build an evolutionary argument for these aspects. We conclude that there is strong evidence that much of human morality has evolved and a long (but possible) way to go to build evidence for how a uniquely human morality might have evolved. One offshoot of this grounding of morality in evolution is that it highlights the intimate links between the development and expression of morality and immorality, a theme that will be woven throughout the book.

1.2 Basic Evolutionary Processes (With an Eye Toward Morality)

“Survival of the fittest” is a phrase that Darwin first uses in his fifth edition of *The Origin of Species* (Darwin, 1869 p. 92). He uses it as a synonym for natural selection. The many misunderstandings of this phrase over time have come to haunt those who would talk about evolution, and particularly evolution and morality. Herbert Spencer originated the “survival of the fittest” phrase (Spencer, 1866, pp. 444–445) and, among other burdens the phrase carries, it has come to be associated with Social Darwinism, a much maligned attempt to derive a kind of “might makes right” ethical principle.³ The phrase has also been misunderstood to suggest that it is the most physically fit, or aggressive, or selfishly competitive individual who survives. This is a mistake on multiple fronts. Natural selection is not about individual survival, but about the ability to reproduce (and thus propagate one’s genome). And since it is about reproduction, it need not be about physical fitness or aggressiveness, but also about cooperation or camouflage or cleverness or compassion that allows the production and continuation of the genome. As we will see later, it is also a mistake in terms of the level at which one can speak of selfishness. Genes may metaphorically be selfish (Dawkins, 1976), but individuals do not have to be (de Waal, 2006).

However, to understand Darwin’s approach to “natural” selection, one must first understand the material on which it works. Darwin begins his *Origin of Species* (Darwin, 1859) not by talking about natural selection but about variation. Any population (of sheep, pigeons, or insect-eating plants) has some natural variation in the characteristics of individuals in that population. Darwin begins his argument by noting that breeders (of pigeons or of strawberries, etc.) take advantage of this variation and selectively breed individuals with characteristics they desire. Darwin calls this “selection by man.” Natural selection is a similar process operating without (usually) human intervention. Some individuals with particular characteristics survive long

³See Weinstein (2009) for a balanced discussion of Spencer’s role in Social Darwinism.

enough to pass on those characteristics to their offspring. To the extent that those characteristics have played a role in the propagation of characteristics in further individuals, then one can say those characteristics have been “naturally” selected or that they are evolutionary “adaptations.” This, then, is the basic idea of selection.

Scholars of evolution list many and various varieties of selection (see Sober, 2006 for an overview of these and other issues). Because the moral is so closely associated with the social, the kind of selection that will particularly concern us here is *social selection*. This is natural selection driven by social interaction within a species, selection of the sort that may have resulted in the rapid increases in brain size mentioned earlier. There are four versions of social selection we will deal with here: sexual, kin, and group selection and gene–culture coevolution.

1.2.1 Sexual Selection

Though it may certainly have other functions, a primary evolutionary function of sexual interaction is the continuation of the species and the genome. In animals, sexual selection occurs during mating choice, with the choice based on some characteristic among the available partners (Krebs, 2008). The classic example in the literature is the mating selection among peacocks, with the female basing her choice in part on the characteristics of the splendid tail of the male.⁴

Can we construct a story of how this might account for moral characteristics among humans? It is at least the case that, across many cultures, moral qualities are preferred in mates (Buss et al., 1990), with dependability and kindness/understanding appearing near the top in lists of preferences for both sexes from thirty-three different countries. In addition, there is evidence that humans can reliably identify moral dispositions in others based on short observations (Fetchenhauer, Groothuis, & Pradel, 2010), making it possible that selection could be based on this identification. Miller (2007) has marshaled the evidence for a theoretical account of sexual selection of virtues and suggested a set of testable theses based on this account. Others (Goetz, Keltner, & Simon-Thomas, 2010) have included sexual selection among the kinds of selection pressure that can support an evolutionary account of the virtue of compassion in both sexes.

Thus, one can construct a case for the evolutionary sexual selection of virtues. We expect this is often overlooked because of the unfortunate connotations of “survival of the fittest.”

1.2.2 Kin (Family) Selection

The idea of kin selection is implicit in the phrase *inclusive fitness* (Krebs, 2008). Inclusive fitness is based on the number of genes an individual can contribute to succeeding generations both directly (by producing offspring) and indirectly (by supporting family relatives who share genetic material). This concept makes it

⁴The empirical story is somewhat more complicated by the search to determine just what it is about the splendid tale that the peahen responds to (Loyau, Petrie, Saint Jalme, & Sorci, 2008). This cannot, of course, be immediately applied to the human species.

6 Taking Moral Action

possible to understand how compassion toward one's immediate and extended relatives might evolve (Goetz et al., 2010) since care for relatives increases one's inclusive fitness. This pattern of closer care for closer relatives has been documented in a wide variety of cultures (Essock-Vitale & McGuire, 1980). If compassion for kin already exists it can be adapted or generalized to compassion for a wider group (Goetz et al., 2010).⁵

1.2.3 Group Selection

One can extend the idea of selection based on one's relatives to selection based on one's group. Wilson and Sober (1994, 1998) revived the idea of group selection and embedded it in a model of multi-level selection. One can think of selection occurring at the level of the gene, the cell, the individual, the kin group, or of larger, arbitrary groups. Group selection was originally proposed by Darwin (1859). It seemed to later researchers (see Richerson & Boyd, 2005 for a review) that selfish individuals in an arbitrary group would take advantage of cooperation, without themselves helping, increasing their fitness with no cost to themselves. It was thought that mathematical models easily predicted that in this case, the selfish defectors would eventually take over in any group (see Thompson, 1998 for some of the complexities in this claim). This "free riding" seemed to make it impossible for cooperative behavior to evolve based on its adaptive value for a group of unrelated individuals. In response to these concerns, Sober and Wilson (1999) identified the conditions that must obtain for group selection to work.

To make group selection viable, variation on the characteristic (e.g. compassion) *between* competing groups would need to be larger than variation *within* each group. Simply put, the groups must differ on the characteristic. Equally important is that *selection pressure between* the groups would also need to be greater than any *within group selection* that might be caused by free riding. Sufficient selection pressure between groups could be produced by things like intergroup conflict, foraging for a common but scarce resource, or the founding of new groups (Sober & Wilson, 1999).

If group selection is going to work to increase some moral attribute that is favorable to the group, it needs to be the case that individuals can recognize the group and establish an attraction to it. This makes immediate sense in terms of kin groups, and such attraction can develop from already existing attractions associated with parenting (Goetz et al., 2010). But how might this attraction establish itself with arbitrary groups? Lewis and Bates (2010) used a twin study and mathematical modeling to provide evidence for genetically based preferences for particular groups (in this case, ethnicity, religion, and race). In addition to these basic preferences for each of the groupings, they also found evidence for a *central affiliation mechanism* (CAM) that

⁵The technical term for this is *exapted*, the selection of an already existing characteristic to serve another purpose. The existing characteristic need not have been naturally selected before, but it might simply have been a necessary byproduct of an already selected characteristic (e.g. a biologically affective experience like compassion for kin that is available to be exapted for non-kin). It might then play a role in group selection to produce a broader characteristic (Goetz et al., 2010).

allows for affiliation or attachment to arbitrary groupings (like a sports team). This preference to be loyal to groups that develop out of local culture might be naturally selected (in a similar way that sexual selection might select virtues). Thus, there is some support for the idea that loyalty to groups of individuals is a characteristic that might be selected for by evolutionary processes (see also Krebs, 2008).

1.2.4 Gene–Culture Coevolution

One more complication must be added to the ways that evolution is socially shaped. Human culture can itself create new selection pressures on individuals in a population, and thus influence genetic characteristics of the population over time. Culture creates an environment that produces selection pressure on the individuals in that culture. One example of this is the widespread change to agriculture. Richerson and Boyd (2010) provide examples of this sort of coevolution with the change in the enzymes that allow adults to digest milk occurring among those populations that had adopted herding of livestock and also the coevolution of enzymes that help with the digestion of plant starch among those adopting agriculture. Gene–culture coevolution is at the center of debates about the evolution of cooperation and altruism,⁶ since this process involves cultural processes like altruistic punishment⁷ that would “deselect” individuals who violate group norms (Boyd et al., 2003; Boyd, Richerson, & Henrich, 2011).

We are only beginning to understand the complexity of the pathways that might lead to the current variety in human culture and genome. Genes do not code for specific behaviors, but they can influence systems that make some behaviors more likely to be expressed or some things easier or harder to learn (see Schmitt & Pilcher, 2004 for examples). The processes of development in the organism that move from a genetic encoding to a phenotypic (body, behavior, etc.) outcome of that encoding (a process called expression) can interact with an organism’s experience in the world. For instance, careful maternal care can support the expression of genes that make the giving of maternal care more likely when the offspring grows up (Goetz et al., 2010).

1.2.5 Evolution vs. Culture

One of the standard critiques of evolutionary accounts of morality (and indeed one of the standard critiques of evolutionary psychology) is that these accounts ignore the overwhelming influence of culture on the complexity of human behavior (McKinnon, 2005). The dichotomous structure of this argument should immediately make one suspicious, since it mirrors many other sometimes overzealous debates in the social sciences like that between nature or nurture, personality or situation, reason or emotion, altruism or selfishness, etc.

⁶See the section on Selfishness for selection for altruism.

⁷This is punishment of an offender of a norm that is costly to the punisher (thus altruistic).

But before looking at an integrative response to this dichotomy, we should note the difficulties that the “culture alone” camp has with explaining some aspects of morality. Two of the most well-documented psychological accounts of moral judgment and action are the social learning (Bandura & Walters, 1963) and cognitive developmental ones (Boyd et al., 2011; Kohlberg, 1963; Turiel, 2006). Both have difficulties explaining some aspects of morality in comparison to evolutionary accounts. Social learning accounts have trouble explaining how moral rules emerged in culture, why people are motivated to preach them, and why children resist induction of moral rules. Cognitive developmental accounts struggle to explain the centrality of affect in moral judgment, the continuing presence of “childish” moral thinking in adults, and the connection (or lack thereof) between moral reasoning and behavior (Krebs, 2008). In addition, both approaches have difficulty in explaining the pattern of data that surround some effects (such as the incest taboo) that have been extensively investigated from an evolutionary perspective (Schmitt & Pilcher, 2004). Finally, both cultural approaches find it hard to explain (or even finding relevant) the extensive list of cultural universals that have moral content (Brown, 1991; Curry, Mullins, & Whitehouse, 2019). This is not an argument that cultural or purely psychological approaches are wrong but simply that they are incomplete, as the evolutionary approach itself is incomplete. This is the reason we try to bring together multiple narratives of moral behavior in this text.

In Section 1.2.4, we covered the process of coevolution. This is the idea that culturally valued, socialized, and transmitted behaviors can influence the evolutionary path of groups, making things like cooperation an evolutionary advantage for a group. This suggests that culture may influence evolutionary paths more than evolutionary paths may influence the expression of culture. The idea of coevolution has emerged from decades of theoretical and empirical confrontation in the evolution–culture debate (Durham, 1991; McElreath, 2010; Richerson et al., 2010).

This understanding of culture–evolution interaction helps to bring some resolution to the voluminous literature on altruism (de Waal, 2008) as we will see in the A Note on Selfishness and Altruism section. And it also helps us understand the many different ways one might find culture–evolution interaction underlying moral judgment and behavior. Some authors (e.g. Mikhail, 2007) have proposed that much of our moral judgment is innate and pre-packaged into independently operating modules.⁸ But it is more likely that the wide range of moral things in our lives is, from the perspective of evolution, a kludge (Stich, 2006) and is constructed from a range of abilities and predispositions (Cushman & Young, 2011; Cushman, Young, & Hauser, 2006; Heyes, 2020). For instance, moral emotions can channel our reactions (e.g. anger against defectors can provoke us to punish them) in ways that help to enforce group norms and make the group more successful (and in ways that pass on this capacity for altruistic punishment). This is an approach adopted by Haidt and Joseph (2004, 2007), who propose sets of moral emotions linked to dimensions of how actions are morally judged.

⁸See Chapter 7 for more detail.

1.2.6 A Note on Selfishness and Altruism

One of the conceptual oddities of the debate in evolution for the past thirty years has been confusion about the word “selfish.” This is in part because Dawkins (1976) published a book using the “selfish gene” as a metaphor for evolution at about the same time that economic models of the rational and self-interested “Homo economicus” were widely discussed. Dawkins’s argument was that natural selection occurs at the level of the gene, and that genes were metaphorically “selfish” when their “strategies” to replicate themselves were analyzed at this level. The idea received wide acceptance in biology and for some time formed the orthodox view of the level at which selection was operating: the level of the gene. More recently, multi-level approaches to selection (see the discussion in Section 1.2.3) have been proposed and favorably evaluated (Wilson & Sober, 1994, 1998).

However, the confluence of Homo economicus and the selfish gene in the 1970s, along with some unfortunate language that Dawkins used toward the end of his influential book, have served to support the idea that because genes are *metaphorically* “selfish,” then human psychological motivations must be selfish too. There are a variety of problems here. First, the idea simply assumes an identity across these levels of explanation. The word selfish is rooted in a description of one kind of human motivation and action, and it works as a metaphor for the point Dawkins wants to make about genes. But to then take that metaphorical use and deduce a psychological motivation or even a pattern of behavior is at best a stretch requiring good evidence. Genetically selfish is simply different from psychologically selfish. And the case for the psychologically selfish Homo economicus is not as strong as it used to be, and may well be suffering the death of a thousand limiting factors (Aktipis & Kurzban, 2004). An additional difficulty is that a simple “always be selfish” strategy is an implausible one in a shifting environment and may be less successful across different environments (Krebs, 2008).

There is clear evidence for a wide range of cooperative, non-prototypically selfish activities in humans and other animals including: mutualism (e.g. group hunting), simple reciprocity (e.g. exchanging meals), cooperation with cooperators (requiring keeping track of “defectors”), “community concern” (de Waal, 2006), and long-term social investment (e.g. care for infants, family, friends). Work in the laboratory (Aktipis & Kurzban, 2004; Van de Vondervoort & Hamlin, 2018) and in the field across many societies (Curry et al., 2019; Henrich et al., 2005) supports these sorts of cooperative motivations in human action against the simple self-interest model. And work by de Waal and colleagues (2006, 2008) shows similar behavior among other species. The renewal of research on cooperative and compassionate behavior across cultures and species seems to be leading to the “extinction” of the selfish Homo economicus model (Aktipis & Kurzban, 2004).

Even though cooperation, nurturance, and compassion are important and natural human motivations, we should not assume a uniformly bright picture. There is a range of motivations and behavioral strategies in our moral interaction that involves selfishness, deception, taking advantage of others, and outright aggression. Indeed, a profound irony is likely built into the group selection mechanism that may be operating to produce cooperation and caretaking. Effective group selection for cooperation *within* the group requires some level of competition *between* groups

(Sober & Wilson, 1999; Wilson & Sober, 1994, 1998). This competition can be created by warfare, but also other more peaceful means, like competition for a scarce resource. In addition, most models of group selection to produce cooperation assume what is called *vicarious punishment* for in-group members, that is, aggression against those who deviate from the cooperative norm of the group. Both of these mechanisms underpin what has been called the dark side of altruism (Graham & Haidt, 2012; Skitka & Mullen, 2002).

Now that we have a sense of the basic terms of how evolutionary processes might work, we will look into ways they might be combined together to produce evidence for how some aspect of morality might have evolved.

1.3 How to Argue That Behavior Was Selected Because It Was Adaptive

Evolutionary psychology has frequently been accused of consisting of too many “just so” stories⁹ about how some psychological characteristic might have been selected in a hypothetical past (Gould & Lewontin, 1979) – but having only that hypothetical possibility and little evidence. Having a “just so” story is a fine place to begin when that story points to where one might collect evidence. Given the relatively recent state of evolutionary psychology, we might expect for evidence about most proposed psychological adaptations to be fairly sparse, but these networks should become denser as more work is done. How might this happen?

Much of evolutionary psychology is an argument for how some trait, or constellation of traits, might have been selected 100,000–300,000 or more years ago. How does one collect evidence and make an empirical argument for something that happened this long ago? In part, one does it by changing the question. For example, if in fact incest avoidance, a common ethical norm, is an evolutionary adaptation, what pattern of evidence would we expect to find today in incest avoidance’s distribution across cultures, its relationship with genetics, its physiology, etc.? To answer this question requires several different kinds of evidence. Schmitt and Pilcher (2004) provide us with a useful map of eight different areas of evidence that are needed to give a convincing, valid case that some item of interest is an evolutionary adaptation:

1. *Theoretical coherence*: The extent to which the adaptation connects to one or more theories or theoretical networks about evolution. This can involve some of the classical theoretical accounts in biology or psychology or game-theoretic or mathematical modeling theory. This is a central piece of evidence in the network: lacking a coherent theoretical explanation means we would not have information about why or how the item is related to an evolutionary history or even the pattern of data we should expect in the other seven categories. In the case of incest, this consists of evidence that genetic problems arise in offspring of incest (Schmitt & Pilcher, 2004).

⁹This is a reference to Rudyard Kipling’s *Just So Stories* (1912) for children. These are fantastic stories of how, for instance, the camel got its hump or the whale got its throat. The narrative logic in the stories is that it is a pleasing tale, and therefore must be true, just so.

2. *Cross-cultural evidence*: If an adaptation occurs with regularity across cultures, or interacts across cultures in interesting ways (e.g. being less prevalent when cultural norms are strong, or occurring in reaction to specified cultural stimuli), then one may have evidence that the characteristic is influenced, at least in part, by some factor in common across all the cultures. There are almost universal prohibitions against incest across cultures (though what counts as incest varies), and there is cross-cultural evidence that being raised with other-sex peers before age six produces a diminishment of attraction to those peers (Schmitt & Pilcher, 2004).
3. *Hunter-gatherer evidence*: There are still cultures today that inhabit environments much like that of our pre-agricultural forebears. This gives us (somewhat tenuous, but nevertheless helpful) evidence of the likely environmental and cultural pressures for selection that may have been operating at the time the theory says the adaptation should have been adaptive. Ethnographies of these cultures show, again, almost universal absence of incest (Schmitt & Pilcher, 2004).
4. *Psychological evidence*: A particular attribute like incest avoidance may follow a developmental path that seems linked with the function our theory says it should have. Or it might relate to other psychological characteristics (or *not* correlate with them), or serve similar functions in modern societies. All these are part of the network of evidence that the characteristic is playing the adaptive role the theory says it should. Being raised with other-sex peers before age six reduces attraction, and this provides developmental psychological evidence (Schmitt & Pilcher, 2004).
5. *Physiological evidence*: If one can find a physiological substrate for the characteristic in question (e.g. oxytocin and empathy: Rodrigues, Saslow, Garcia, John, & Keltner, 2009), then one has established a physical mechanism by which a characteristic might be heritable. The interaction of this physical mechanism with culture and situational stimuli can enrich the evidence network to provide more detail about how the adapted characteristic functions in society. There is evidence for the sense of smell as a mechanism that produces sexual aversion to close kin, thus incest avoidance (Weisfeld, Czilli, Phillips, Gall, & Lichtman, 2003).
6. *Medical evidence*: This is similar to the physiological evidence but based in medical cases of the supposed physiological mechanism going wrong. When the physiology is upset, are the functional relationships of the characteristics also upset (see e.g. the use of brain damage cases in Chapter 2)? Schmitt and Pilcher (2004) do not list any medical evidence for incest avoidance.
7. *Phylogenetic evidence*: Finding similar patterns in other species, where the theory leads one to expect the patterns, is another set of nodes in the network of evidence. For some characteristics (e.g. the effects of oxytocin on social behavior) one might expect to see a range of species across all mammals (Rodrigues et al., 2009), while for other characteristics (like a sense of fairness) one may expect only similarities among other primates (de Waal, 2006; Preston & de Waal, 2002). In terms of incest, animals rarely mate with parents or siblings (Schmitt & Pilcher, 2004).
8. *Genetic evidence*: We are increasingly able to use genomic evidence to trace the timing of events in human evolution (Richerson et al., 2010) and this will become a rich source of data that can connect with the others in this list. Schmitt and Pilcher (2004) do not list any genetic evidence for incest avoidance.

Schmitt and Pilcher (2004) use this framework to survey the evidence for four psychologically relevant characteristics as evolutionary adaptations. Profet (1988) has followed the possibility that pregnancy sickness and its associated food aversions are an adaptation that allows for the fetus to be protected against toxins in plants that would cause birth defects. This story suggests that pregnancy sickness will follow a particular pattern (certain food avoidances, a particular timeline, the relationship of sickness to miscarriage, similarity across cultures, etc.) and there is much evidence for this rich network of relationships (Schmitt & Pilcher, 2004). Schmitt and Pilcher (2004) find less extensive, but still “exemplary,” evidence for incest avoidance. They find only “extensive” evidence for men’s greater desire for short-term sexual variety than that of women, and “moderate” levels of evidence that we have evolved a mechanism that makes it easier to learn fear of snakes.

1.4 Evolutionary Building Blocks for a Robust Morality

So, what can we learn about our moral nature by seeing it through the lens of evolutionary adaptation? At the most complex level, this might mean we have inherited a distinct and structured universal moral grammar, much like what has been claimed for the deep structure of a universal linguistic grammar (Huebner, Dwyer, & Hauser, 2009; Mikhail, 2007). We will evaluate this claim in a moment, but the picture of morality we provide in this book attempts a much wider claim. Moral action certainly involves a kind of structure of moral intuition, but the full picture includes far more: emotional intuitions, strategic planning, self-regulation, self-concept and group identity, and much more. Each of these might have an evolutionary underpinning. But it seems unlikely that they would have all evolved together as one piece. The usual pattern of evolutionary adaptation that moves toward complexity is more one of building that complexity slowly based on some existing variation.¹⁰ Given that many of these aspects, such as opposing values, are often at odds with one another (and that self-regulation is, in one sense, a mechanism to help adjudicate these conflicts), an evolutionary story would expect the complexity of moral action in humans to have been built up slowly across many species and a great deal of time. For instance, maternal care for the young is common to all mammals, but it plays an important part in human morality (Preston & de Waal, 2002).

For an overview of how these building blocks might come together, we will follow the proposal by Krebs (2008) of the various psychological building blocks that likely came together to constitute what we now call morality. This overview itself will suggest additional pieces that remain to be covered, and we address them in Section 1.5.

1.4.1 Adaptive Sociality

The first and fundamental building block that Krebs lists is that of sociality – or living in groups. Group living is built out of interdependence among animals and the

¹⁰Evolution does not always build toward complexity. Even “build” is a metaphor that contains intention. “Natural” selection has no intention. An organism may evolve toward less complexity if that makes it fit better in an environment.

complexities of how that interdependence is negotiated. This sociality is the basic background for the development of all the other building blocks. One thing to note: in Krebs's phrase, social life is built of both confluence *and* conflict of interest. For organisms to work in a group there must be some attachment to the group, some confluence of interest, and this attachment can be elaborated in many directions. As de Waal (2002) and Goetz et al. (2010) have documented, one origin (at least in mammals) of this social attraction must have been the giving and receiving of care between parents and offspring, which became more crucial as the time for complete maturation of offspring increased. And there is of course conflict of interest that must also be managed. This conflict and confluence is the arena in which moral action and judgment evolved.

1.4.2 Pro-social Behavior Strategies Based in Social Emotions

Emotions are not simply “feelings” but are instead a complexity of: (1) *cognitive appraisal* (that arouses and shapes the emotion); (2) *directed arousal* (that provides the “feel” of the emotion and signals its urgency); and (3) *action tendency* (that links action to the reaction of the emotion). (See Oatley, Keltner, & Jenkins, 2006 or Chapter 8 for an overview). How this complex itself has evolved is a fine puzzle; but for social animals, it certainly evolved in connection with socially relevant behavior. Krebs (2008) lists three relevant sets of emotions tied to pro-social behaviors: (1) *deference* in dominance hierarchies based in fear and respect; (2) *cooperation* among individuals based in gratitude, anger, and indignation; and (3) *altruism* based in love, sympathy, and empathy. We suggest some more in Chapter 8.

1.4.3 Strategic Interaction

Simply having social emotions and reacting based on them produces a sort of mechanical social order. But, once sociality is established, cheating is possible when one can take the other's perspective in order to deceive. A basic requirement for strategic social interaction is *theory of mind*, recognizing and ascribing mental states to ourselves and others (Wellman, 2018). There is evidence that even infants have basic aspects of theory of mind (Hudac & Sommerville, 2020) though its developmental trajectory looks similar to that of reading (Heyes, 2020), suggesting that the full capacity is some mix of built-in capability and cultural shaping.

The capacity to take perspective can then produce an “arms race” (Krebs, 2008) when it makes it possible to detect cheating and punish free riders. It can become even more complex when we keep track not only of cheaters but also those among us who do not punish the cheaters. (Boyd, Gintis, & Bowles, 2010; Boyd et al., 2003). This strategic level of interaction can apply to areas as abstract as what de Waal has called “community concern,” the strategic intervention in group processes to produce desired outcomes at the group level (de Waal, 2006).

1.4.4 Conscience

Krebs (2008), citing Darwin (1859), lists the linking of emotional responses with anticipatory fear of punishment from others (and later the internalized other) as a potential origin for what we call conscience. There is some evidence for this linkage.

For instance, physical responses associated with negative self-emotions (like shame and guilt) mirror behavior shown in appeasement displays in humans and other primates (Keltner & Buswell, 1997; Leary, 2007). Leary (2007) provides an overview of the research that makes this link between what we might call conscience and anticipatory reactions of real or imagined others. But one need not assume that this effect runs on anticipatory punishment alone. One can find evidence of an evolved capacity for compassion that can motivate behavior, not merely of avoiding the bad but also approaching the good (Goetz et al., 2010).

1.4.5 Moral Judgment and Reasoning

Krebs (2008) proposes that the origin of moral reasoning lies in moral communication – in verbal influence attempts to get others to behave in morally prescribed ways. But this is surely starting too late in the process. Tse (2008) looks back even further to a proposed delinking of attention from concrete objects in the world that allows arbitrary symbols to be associated with each other.¹¹ This is likely a basis for the flexibility of human reasoning, and the basis of things like metaphor, allowing for the association of arbitrary symbols and ideas in metaphor (a poem as lovely as a tree) but also the association of actions with abstract classes of good and bad. There is currently a vigorous debate about whether something as structured as a moral “grammar” is an evolutionary heritage of humans. This idea is an analogy from what seems to be an evolved grammar in language. There is currently only suggestive evidence for this hypothesis (Beller, 2010) and nothing like the strong networks of evidence we have reviewed for other adaptations (see Greene, 2008; Mikhail, 2007 for some idea of the controversy).

1.4.6 Moral Norms

Earlier we reviewed at least one moral norm (the incest taboo) that seems to have an evolutionary basis (Schmitt & Pilcher, 2004). But we clearly have a large store of moral norms that are in regular use, and that vary widely from culture to culture. In Chapters 7 and 8, we review the evidence that there is a structure to these norms that is cross-culturally consistent, with the same value sets showing up across all cultures but differently emphasized by each culture (Haidt & Joseph, 2004; Haidt, Koller, & Dias, 1993; Schwartz, 2006). This cross-cultural evidence of a universal structure might argue for an evolutionary origin (see Haidt & Joseph, 2007 for an evolutionary argument). Managing these norms involves a complex cultural process of protecting against deviance while promoting diversity (Wright, 2021).

We already have a long list of building blocks, and Krebs (2008) provides some idea of how these might come together to support what we call moral judgment and action. However, the chapters in this text suggest we need to add even more complexity to this picture. There must be some accounting of the development and

¹¹There is currently little evidence for this interesting proposal. Whether one calls this an interesting theoretical proposal or a “just so story” depends on one’s attitude toward the likelihood that such evidence will turn up (Dietrich, 2008).

integration of short- and long-term self-regulation. How and why might willpower (and variations in willpower) have evolved (Heatherton, 2011)? While it is merely difficult to assemble the network of evidence for short-term self-regulation as an evolutionary adaptation, work on the evolutionary underpinnings of longer-term self-regulation seems quite thin (Heatherton, 2011); though Miller (2007) suggests the sexual selection of moral virtues like dependability might play a role.

We know that storytelling (McAdams, 2009) and religion (Graham & Haidt, 2010; Kirkpatrick, 1999) play a role in moral judgment and action. For many of these more complex and higher-level aspects of morality it may well be that the relatively new approaches to culture and coevolution (Wilson & Sober, 1998) will provide us with a way to understand their construction. We will now look at a list of many of these additional aspects of morality that may (or may not) find some evolutionary underpinning.

1.5 Human Distinctiveness: How Large is the Gap?

We have spent most of this chapter looking at the processes of evolution and the building blocks of human morality. The goal has been to see how we can ground human morality in, or derive it from, various aspects of the capacities and behavior of our ancestors. An older version of this approach, disparagingly called “vener theory” by de Waal (2006), proposes that humans have only a veneer of politeness and hypocrisy overlaying the motivation of brute selfishness. Higgins and Pittman (2008) suggest that this version of human moral evolution is mirrored in the historical split in evolutionary approaches in different sub-fields of psychology. In cognition, we are said to have “developed from” prior species, while in motivation, we have been presented as merely “deriving” our animal motivations from prior species. This results in an image, they claim, of “humans as having the mind of a god and the motives of a brute” (Higgins & Pittman, 2008, p. 363). In the realm of morality, this image of the supremacy of reason is aided and abetted by philosophical approaches to ethics (e.g. Kant, 1999; Rawls, 1971/1999) that portray reason as both the foundation and the limit of real (that is, human) morality.

But current approaches to human morality that we review in this book suggest that: (1) human morality is constituted of more complex stuff than pure reason; and (2) the differences with other animals and our primate ancestors lie in a *range* of distinctly human abilities and motivations (Higgins & Pittman, 2008) and not just differences in reason alone. There is unanimity in recognizing that there are deep differences between humans and other animals in morality but the nature, structure, and meaning of those differences is still in dispute.¹²

¹²Some claims might underestimate these differences (Adolphs, 2009) in the way veneer theory does. Other claims may underestimate the unique abilities of other animals by making human morality the criterion, and missing uniquely adapted, species-specific capacities of other animals (Emery & Clayton, 2009). Crows, for example, do not have a full-blown theory of mind like humans, but they are remarkably good at hiding food from other crows (Emery & Clayton, 2009). De Waal (2006, 2009), Emery and Clayton (2009), and Adolphs (2009) provide some insight into the debates about these differences.

The dimensions associated with morality on which humans differ from other animals include at least the eleven categories noted in Sections 1.5.1–1.5.11. We have constructed this incomplete list from the literature in evolutionary psychology. The effort to build coherent networks of evidence for the origins of human ability in each of these categories may, of course, suggest additional categories or the merging of some. The list makes it clear that we are, indeed, at the very beginning of an evolutionary understanding of the structure and development of a truly human morality.

1.5.1 Symbolizing and Reasoning

The beginning of the ability to reason lies, perhaps, in the ability to take any particular representation of an object and link it arbitrarily to some other mental representation, e.g. a poem is like a tree (Tse, 2008). This may provide the basis for human ability to represent things symbolically (think of early cave paintings and whether any animal other than humans would “get” them). The empirical evidence for an evolutionary basis for moral reasoning and judgment remains unclear (Beller, 2010). Some researchers claim that regularities in the way people ignore logical structure and reason in specific content areas (e.g. cheating, incest) show the primacy of evolutionary modules for content-specific reasoning (Cosmides & Tooby, 2008; Cushman et al., 2006).¹³ But this claim is disputed by both philosophers (Prinz, 2008) and psychologists (Chater & Oaksford, 1996; Pietraszewski & Wertz, 2021). This is a disagreement about *how* ought-based (or deontic) reasoning might have evolved in humans. But it is also a dispute anchored in an agreement that human capabilities for deontic reasoning are vastly more complex than that of other animals.

1.5.2 Time Perspective

Time perspective underlies the ability to think about the moral consequences of an action or to make plans to achieve moral goals. In order to support some conception of a past or a future time, or of an alternative present, one must construct a mental representation and distinguish it from one’s own current experience (Higgins & Pittman, 2008). Thus, a conception of time requires the symbolizing abilities mentioned in Section 1.5.1. It remains unclear whether any animals other than humans can experience an imagined point of view (Adolphs, 2009) and thus have some sense of time. The sense of the future and past allows humans to construct moral stories about the past and about the future, and compare the self to them (Higgins & Pittman, 2008). This adoption of another point of view than that of immediate consciousness is likely one that underlies the ability to recognize other minds (Adolphs, 2009).

¹³For an example of this content-specific reasoning, see the discussion on moral dumbfounding in Chapter 7. In short, careful experimentation can find instances where people will simply say “I know it is wrong, but I do not know why.” Such moral intuition suggests that these particular moral judgments do not derive from a general reasoning mechanism. But this does not mean that no moral judgments derive from a general reasoning mechanism.

1.5.3 Cognitive Control and Self-Regulation

As we argue in Chapters 6 and 9, self-regulation is central to moral action. It is certainly the case that other animals can exert cognitive control over immediate desire, both to deceive other animals (de Waal, 2006) and in response to training by humans (H. C. Miller et al., 2010). This sort of immediate willpower of response suppression seems to share similar biological substrates in humans and other mammals (e.g. it uses glucose-based energy: Baumeister, Bratslavsky, Muraven, & Tice, 1998; H. C. Miller et al., 2010). But the extent, complexity, and flexibility of human self-control mechanisms far outpaces those of other animals (Heatherston, 2011). It extends to what the philosopher Frankfurt (1971) has called “second order desires,” or the desire to have a particular desire and to the planning of outside influences that will help one attain future states (for example, by going to school: Cohen, 2005). It also requires directed planning and practice to achieve long-term goals (Baumeister, Masicampo, & Vohs, 2011; Cervone, Shadel, Smith, & Fiori, 2006; Lord, Diefendorff, Schmidt, & Hall, 2009; McRae, Ochsner, & Gross, 2011).

1.5.4 Extensive Integration of Reason and Emotion

Moll and colleagues (de Oliveira-Souza, Zahn, & Moll, 2016; Moll, de Oliveira-Souza, & Eslinger, 2003; Moll, De Oliveira-Souza, & Zahn, 2009) have hypothesized that the most important thing about human reason may well be its *integration* with rather than its opposition to emotion.¹⁴ Even philosophical champions of reason as a marker of human morality (Korsgaard, 2006) have suggested that the integration of reason and emotion is a central puzzle in the evolution of morality (Korsgaard, 2010). It is not simply the *addition* of rationality to an “emotional animal” that makes humans special and human morality possible. It is the extensive integration of that rationality with emotion that allows them to influence and “tune” each other. Emotions influence our implicit and explicit moral reasoning by guiding intuition (Haidt, 2001) and reasoning (Goldin et al., 2008; McRae et al., 2011). And rationality influences emotion through the multifaceted self-regulation processes mentioned in Section 1.5.3, through the appraisal and re-appraisal process (Lazarus, 1991; Ochsner et al., 2010), and through the planned construction of culture (Cohen, 2005).

1.5.5 Extensive Culture

Humans have complex, stable, and constantly adapting cultural systems with enduring artifacts, language, artistic expression, economics, ritual, values, social structures, and extensive bodies of knowledge and practice (McElreath, 2010; Richerson & Boyd, 2005). These cultural achievements are clearly one of the major differences between humans and other animals. There is some evidence for social transmission of information among other primates (Adolphs, 2009) and thus some rudimentary socially shared culture. Understanding the linkages among evolution, neuroscience, and cultural processes will be a crucial undertaking for a complete understanding of moral action (Kitayama, Varnum, & Salvador, 2019; Wright, 2021).

¹⁴We review evidence for this integration in Chapters 2 and 8.

1.5.6 Self-Consciousness and Self-Awareness

The classic experimental demonstration of self-consciousness in children and in non-humans is the body awareness test. This involves placing some marker on the body in a place that is difficult to see without a mirror and then measuring the attempts to examine the new decoration. Evidence for such awareness has been found in chimpanzees, bonobos, dolphins, elephants, and monkeys (Adolphs, 2009). This seems adequate to demonstrate body awareness but its evidential status for awareness of one's own mind is unclear (Adolphs, 2009; Gallup & Anderson, 2020). For normally functioning adolescent and adult humans, there is unanimity that a complex, multidimensional, and morally relevant self-awareness is present (Byrne & Whiten, 1988; Robins, 2021).

1.5.7 Social Consciousness

Dogs and some other domesticated animals, perhaps because they have been shaped by human intervention over many generations, are skilled at interpreting gestures (e.g. pointing) and in responding correctly in the guesser–knower paradigm (in which they must choose a person based on that person's knowledge of where some food is hidden) (Emery & Clayton, 2009). But dogs do not appear to have some basic theory of mind characteristics that human infants share, like preferring helpers over hinderers in an attention task (Hamlin, 2013; McAuliffe et al., 2019; Van de Vondervoort & Hamlin, 2018). Primates are particularly skilled at tracking kinship and dominance relationships over time (Adolphs, 2009). However, if one asks for a full-fledged human theory of mind in other primates, it appears to be lacking (Emery & Clayton, 2009) – although one can likely find components of it. For example, there is evidence that other primates know that others know something but probably cannot grasp the complexity that the other's knowledge may be false (Emery & Clayton, 2009). Waytz and colleagues (Waytz, Cacioppo, & Epley, 2010; Waytz et al., 2010) have shown that humans' perception of agency (even the anthropomorphic perception of agency in inanimate objects) includes a moral dimension allowing for blame and suggesting that the roots of moral perception go quite deep into basic perception. And though other primates form coalitions (de Waal, 2007) the complexity of human tribalism is much greater (Clark et al., 2019).

1.5.8 Manipulation of Others' Perception

There is evidence that other primates understand and practice deception (Adolphs, 2009; Emery & Clayton, 2009). There is even evidence that crows and other corvids cache food in ways that make it difficult for other crows to find it, but only if other crows who might take it are present (Emery & Clayton, 2009). But again, even in young children, human manipulation of others' perception is vastly more complex, multidimensional, and subtle (Higgins & Pittman, 2008). Humans are deeply concerned about what others think (Higgins & Pittman, 2008) and particularly what others think about them (Leary, 2007).

1.5.9 Sharing Reality With Other People

Humans have a deep need to share their understanding of the world with others (Higgins & Pittman, 2008). All mammals have some abilities at social coordination and signaling (Goetz et al., 2010) and some have components of a theory of mind (Emery & Clayton, 2009). But even human children are competent in communicating shared social knowledge, sharing joint attention, and explicit teaching. Language is a primary achievement of this need (Byrne & Whiten, 1988).

1.5.10 Universalized Normative Evaluations

People's concern about what others think of them has been included in models of the "generalized other" (Heatherton, 2011; Mead & Morris, 1934) as one source of how humans universalize normative evaluations. The philosopher Korsgaard (2006) argues that this concern is more than simply worrying about how others might think of us, it involves seeing a moral reason as a reason for behavior, and using that reason to guide planning and action. In short, it involves some version of moral autonomy where the desires associated with moral reason become dominant over other desires or appetites.¹⁵ Managing these norms in a society involves a complex cultural process of protecting against deviance while promoting diversity (Wright, 2021).

Psychologically, these reasons are experienced as "natural law," as a moral imperative that trumps any social convention or personal preference (Shweder, 2012; Shweder, Mahapatra, & Miller, 1987); and this experience is one of the aspects of human morality that is a cultural universal. Despite clear evidence that other primates have some expectation about the nature of their social world, and show clear distress and attempts at punishment or repair when they are violated, it seems unlikely that one would find evidence for this aspect of moral expectation in non-human primates (Kitcher, 2006; Korsgaard, 2006). Indeed, it is unclear how one could even collect such evidence.

1.5.11 Self-Actualization, Religious Experience, and Self-Transcendence

Although some evolutionary theorists would prefer to dismiss self-actualization since it is not "clearly linked to reproductive fitness" (Schaller et al., 2010, p. 336), others find it both compatible with evolutionary theory and essential in understanding human experience (Peterson & Park, 2010). Both sides in the controversy agree that they do not expect to find it in other animals. A similar pattern emerges with another human universal: religion (Brown, 1991). Though there is disagreement about the status of religion and its relationship to evolutionary theory (Dawkins, 2006; Kirkpatrick, 1999), any suggestion that one might find something like human religion in other animals is conspicuous by its absence. Self-transcendence goes beyond the fulfillment of the self, and invests the self in some purpose beyond the self (Maslow, 1969). It seems to be both a natural product of adult development and a

¹⁵Note how desire and reason are integrated in this account.

process that can lead to extraordinary moral commitment (Walker & Frimer, 2015). This commitment has a dark side, depending in part on the cause one chooses (Koltko-Rivera, 2006; Skitka & Mullen, 2002).

1.6 Discussion

1.6.1 Conclusion

An evolutionary understanding of morality can primarily help us in understanding our capacities for taking moral action. How those capacities are shaped, influenced, and directed is the subject of the other chapters of this book. However, an understanding of capacity, its structures, and its limitations can help to articulate the structure and limits of our morality, help us see how good and evil are linked, and allow us to grasp the precarious human condition that is balanced in the gray zone between those two:

1. *Moral action is deeply tied to the social nature of being human.*

Much philosophy and psychology have presented morality as the achievement of reason and will alone. Evolutionary theory confronts us with how our human morality is deeply rooted in the social nature of our species. The need to negotiate a complex social hierarchy within groups and to compete between groups has produced a set of capacities that we list in Section 1.4. These capacities are densely interwoven and help us to see how moral action results from the complex interplay among social emotions, planning, and rationality. Any understanding of morality that does not take this interplay into account will be inadequate.

2. *The dark side of morality is, ironically, enabled by our desire to do good.*

An evolutionary approach to morality also helps us to understand the peculiar dual-sided interaction of morality with social motives and structures. Social status in social hierarchies and in between groups is based in respect but also anger and fear. Cooperation among individuals and groups is based in gratitude and respect but also vicarious anger and indignation. Compassion for others is based in love and empathy but also vicarious anger and revenge against those who hurt loved ones.

This complex set of social emotions can be harnessed by our rationality and culture in the service of a range of goals, producing both international aid organizations and organized ethnic cleansing and self-giving care for a partner and abuse for a partner. In this way, the psychology of being moral is also the psychology of evil.

1.6.2 Application

In the Introduction, we presented the story of Rick Munson and his wife.¹⁶ The Munsons began by hosting foster children, then went on to host disabled foster children. Eventually, they founded a charitable organization and a facility that provided

¹⁶Rick Munson is a pseudonym for a participant in a study of volunteerism (Hart & Atkins, 2006). The report of the study does not give a pseudonym for his wife, though it makes it clear they made decisions together.

integrated services to children with physical challenges. One can chart this as a move from the pro-social integration of the two individuals into a committed couple, followed by the expression of their concern for others in caring for children, followed by the extension of that concern to community action. Their taking in of foster children already stretches our conception of a narrow evolutionary basis for pro-social action, as it was likely an extension of the desire to have children to those who were clearly outside their genetic family. But their actions were surely due to a capacity for compassion that is a part of their evolutionary heritage and was shaped and supported by their religious culture. It was also made possible by an organized cultural practice of finding homes for children without any. The empathy the Munsons had for other disabled children and their parents and the couple's capacity for planning and rationality led them to a more general community concern, which was expressed by volunteering in charitable organizations. This in turn led to their founding of organizations to meet community needs.

One can see in this example the integration of evolutionarily shaped emotional and rational capacities with culturally constructed social structures to produce good. This provides an example of the fit of these capacities (rationality, planning, compassion, empathy) to the local ecology in which the Munsons were embedded. But it is also an example of the shaping of those capacities by the wider moral ecology (e.g. religious practice) and the local ecology (the lack of integrated service for disabled children and the practice of founding organizations to meet needs). An understanding of the evolutionary basis of morality will require this sort of complexity.

1.6.3 Open Questions

1. *We have an excellent beginning of a model of the evolution of morality and a long way to go in developing it.*

The review of building blocks (Section 1.4) suggests that we are beginning to accumulate an evidential network of the underpinning of the moral ability in humans. But a comparison of the standards for making an evolutionary argument (Section 1.3) with the evidence for each building block suggests we still have a long way to go to fill in the argument.

2. *There is a significant collection of capacities that underlie moral action for which we so far have no viable evolutionary explanations.*

Section 1.5 gives a list of eleven areas in which there is an easily discernable distinction between the moral capacities of humans and other animals. There is clearly a vast field of opportunity for good research and theory here. But doing this work will require changing the way the question is asked. Instead of asking *whether* humans are different from other animals in their morality, we should begin asking *how* they are different, and how those differences have come about. It is some combination of these morally relevant differences in *how* that differentiates human morality from what might be called the proto-morality of other species.

3. *Exploring the fit of the evolved moral capacities of humans with their social and physical ecology will provide for a more complete evolutionary understanding of human morality.*

A central mistake in work on theory of mind is the tendency to measure other animals by their differences from human theory of mind, rather than by the

ecologically valid fit of the abilities of the particular species with that species' social and physical ecology. The field of moral psychology should take a clue from this mistake and insist on an ecologically valid explanation of human morality. Thus, rather than simply looking for universals, we should look for a more complete understanding of human morality in the fit of the moral capacities of humans with their varied social and physical ecologies. This will necessarily include an evolutionary perspective, but is unlikely to be reducible to only that perspective.

4. *The evolution of morality and immorality are closely linked and interdependent.*

Neither morality nor immorality is more "basic." Aggression against outsiders and norm-violators is integral to our understanding of how moral commitment to a group evolved. We can regulate this, and even balance it with compassion, but altruistic punishment based in revenge motives will likely remain central to doing so. Empirical and theoretical work might help us find ways to motivate pro-social behavior and discourage anti-social behavior that work within the limits of our evolutionary heritage and are still acceptable to ethical norms.

1.7 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Aktipis and Kurzban (2004). "Is Homo economicus extinct? Vernon Smith, Daniel Kahneman and the Evolutionary Perspective." An historical overview of the idea of the selfish "economic human" and its diminishing psychological appeal in the face of a wave of research suggesting its falsity (or at least incompleteness).
- Curry et al. (2019). "Is it good to cooperate?" An excellent example of cross-cultural work to establish an evolutionary basis for morality.
- de Waal (2008). "Putting the altruism back into altruism: The evolution of empathy." Primatologist Franz de Waal provides a useful presentation of the evidence for the evolution of empathy and its role in altruism.
- Gallup and Anderson (2020). "Self-Recognition in animals." A review of a half-century of research on animal consciousness that argues that the gap between great apes/humans and other species is larger than usually acknowledged.
- Krebs (2008). "Morality: An evolutionary account." A thoughtful presentation of the main issues in an evolutionary account of human morality.
- Schmitt and Pilcher (2004). "Evaluating evidence of psychological adaptation: How do we know one when we see one?" A readable and systematic overview of what it means to say that some psychological characteristic is an evolutionary adaptation.
- Sober (2006) *Conceptual Issues in Evolutionary Biology* (3rd ed.). Excellent volume edited by a philosopher deeply involved in the evolutionary debates within the biological disciplines. Every chapter is an example of how to think carefully about evolution. Essays on what fitness means, levels of selection, race, culture, ethics, etc.

References

- Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology*, 60, 693–716.
- Aktipis, C. A., & Kurzban, R. O. (2004). Is Homo economicus extinct? Vernon Smith, Daniel Kahneman and the Evolutionary Perspective. *Evolutionary Psychology and Economic Theory*, 7, 135–153.
- Aristotle. (1941). Politics (W. D. Ross, Trans.). In R. McKeon (Ed.), *The Basic Works of Aristotle* (pp. 1127–1324). New York: Random House.
- Bandura, A., & Walters, R. H. (1963). *Social Learning and Personality Development*. New York: Holt, Rinehart and Winston.
- Baumeister, R. F., Bratslavsky, E., Muraven, M. et al. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.
- Baumeister, R. F., Masicampo, E. J., & Vohs, K. D. (2011). Do conscious thoughts cause behavior? *Annual Review of Psychology*, 62(1), 331–361. doi:10.1146/annurev.psych.093008.131126
- Beller, S. (2010). Deontic reasoning reviewed: psychological questions, empirical findings, and current theories. *Cognitive Processing*, 11(2), 123–132.
- Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328(5978), 617–620. doi:10.1126/science.1183665
- Boyd, R., Gintis, H., Bowles, S. et al. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6), 3531–3535. doi:10.1073/pnas.0630443100
- Boyd, R., Richerson, P. J., & Henrich, J. (2011). Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioral Ecology and Sociobiology*, 65, 431–444. doi:10.1007/s00265-010-1100-3
- Brown, D. E. (1991). *Human universals*. Philadelphia: Temple University Press.
- Buss, D. M., Abbott, M., Angleitner, A. et al. (1990). International preferences in selecting mates: A study of 37 cultures. *Journal of Cross-Cultural Psychology*, 21(1), 5–47. doi:10.1177/0022022190211001
- Byrne, R. W., & Whiten, A. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. New York: Oxford University Press.
- Cervone, D., Shadel, W. G., Smith, R. E. et al. (2006). Self-regulation: Reminders and suggestions from personality science. *Applied Psychology*, 55(3), 333–385.
- Chater, N., & Oaksford, M. (1996). Deontic reasoning, modules and innateness: A second look. *Mind & Language*, 11(2), 191–202.
- Clark, C. J., Liu, B. S., Winegard, B. M. et al. (2019). Tribalism is human nature. *Current Directions in Psychological Science*, 28(6), 587–592. doi:10.1177/0963721419862289
- Cohen, J. D. (2005). The vulcanization of the human brain: A neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives*, 19(4), 3–24.
- Cosmides, L., & Tooby, J. (2008). Can a general deontic logic capture the facts of human moral reasoning? How the mind interprets social exchange rules and detects cheaters. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 1, pp. 53–119). Cambridge, MA: MIT Press.
- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1), 47–69. doi:10.1086/701478
- Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from non-moral psychological representations. *Cognitive Science*, 35(6), 1052–1075.

- Cushman, F., Young, L., & Hauser, M. D. (2006). The role of conscious reasoning and intuition in moral judgment. *Psychological science*, 17(12), 1082–1089.
- Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (1st ed.). London: John Murray.
- Darwin, C. (1869). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (5th ed.). London: John Murray.
- Dawkins, R. (1976). *The Selfish Gene*. New York: Oxford University Press.
- Dawkins, R. (2006). *The God Delusion*. New York: Houghton Mifflin Harcourt.
- de Oliveira-Souza, R., Zahn, R., & Moll, J. (2016). The Neuropsychiatry of moral cognition and social conduct. In S. M. Liao (Ed.), *Moral Brains: The Neuroscience of Morality* (pp. 203–236). New York: Oxford University Press.
- de Waal, F. B. M. (2006). Morally evolved: Primate social instincts, human morality, and the rise and fall of “vneer theory.” In F. B. M. de Waal, S. Macedo, & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 1–58). Princeton, NJ: Princeton University Press.
- de Waal, F. B. M. (2007). *Chimpanzee Politics: Power and Sex among Apes*. Baltimore, MD: Johns Hopkins University.
- de Waal, F. B. M. (2008). Putting the Altruism Back into Altruism: The Evolution of Empathy. *Annual Review of Psychology*, 59(1), 279–300. doi:10.1146/annurev.psych.59.103006.093625
- de Waal, F. B. M. (2009). Darwin’s last laugh. *Nature*, 460(7252), 175.
- Dietrich, M. R. (2008). A Just-so story for symbolic thought? Comment on Tse. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Evolution of Morality: Adaptations and Innateness* (vol. 1, pp. 299–301). Cambridge, MA: MIT Press.
- Durham, W. H. (1991). *Coevolution: Genes, Culture, and Human Diversity*. Stanford, CA: Stanford University Press.
- Emery, N. J., & Clayton, N. S. (2009). Comparative Social Cognition. *Annual Review of Psychology*, 60, 87–113.
- Essock-Vitale, S. M., & McGuire, M. T. (1980). Predictions derived from the theories of kin selection and reciprocity assessed by anthropological data. *Ethology and Sociobiology*, 1(3), 233–243.
- Fetchenhauer, D., Groothuis, T., & Pradel, J. (2010). Not only states but traits – Humans can identify permanent altruistic dispositions in 20 s. *Evolution and Human Behavior*, 31(2), 80–86. doi:10.1016/j.evolhumbehav.2009.06.009
- Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68(1), 5–20.
- Gallup, G. G., & Anderson, J. R. (2020). Self-recognition in animals: Where do we stand 50 years later? Lessons from cleaner wrasse and other species. *Psychology of Consciousness: Theory, Research, and Practice*, 7(1), 46–58. doi:10.1037/cns0000206
- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin*, 136(3), 351–374. doi:10.1037/a0018807
- Goldin, P. R., McRae, K., Ramel, W. et al. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, 63(6), 577–586.
- Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205(1161), 581–598.
- Graham, J., & Haidt, J. (2010). Beyond Beliefs: Religions Bind Individuals into Moral Communities. *Personality and Social Psychology Review*.
- Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. R. Shaver & M. Mikulincer (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil* (pp. 11–31). New York: American Psychological Association Books.

- Greene, J. D. (2008). The secret joke of Kant's soul In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Neuroscience of Morality: Emotion, Brain Disorders, and Development* (pp. 35–79). Cambridge, MA: MIT Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.
- Haidt, J., & Joseph, C. (2007). The moral mind: How 5 sets of innate moral intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind* (vol. 3, pp. 367–391). New York: Oxford.
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, 65, 613–613.
- Hamlin, J. K. (2013). Moral Judgment and Action in Preverbal Infants and Toddlers. *Current Directions in Psychological Science*, 22(3), 186–193. doi:10.1177/0963721412470687
- Hamlin, J. K. (2018). Social behavior: Bonobos are nice but prefer mean guys. *Current Biology*, 28(4), R164–R166. doi:10.1016/j.cub.2017.12.054
- Heatherton, T. F. (2011). Neuroscience of Self and Self-Regulation. *Annual Review of Psychology*, 62(1), 363–390. doi:10.1146/annurev.psych.121208.131616
- Henrich, J., Boyd, R., Bowles, S. et al. et al. (2005). “Economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(06), 795–815.
- Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, 3, article 31. doi:10.3389/neuro.09.031.2009.
- Heyes, C. (2020). Psychological mechanisms forged by cultural evolution. *Current Directions in Psychological Science*, 29(4), 399–404. doi:10.1177/0963721420917736
- Higgins, E. T., & Pittman, T. S. (2008). Motives of the Human Animal: Comprehending, Managing, and Sharing Inner States. *Annual Review of Psychology*, 59(1), 361–385. doi:10.1146/annurev.psych.59.103006.093726
- Hudac, C. M., & Sommerville, J. A. (2020). Understanding other's minds and morals: Progress and innovation of infant electrophysiology. In J. Decety (Ed.), *The Social Brain: A Developmental Perspective* (pp. 179–197). Cambridge, MA: MIT Press.
- Huebner, B., Dwyer, S., & Hauser, M. D. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences*, 13(1), 1–6. doi:10.1016/j.tics.2008.09.006
- Kant, I. (1999). *Critique of Pure Reason* (P. Guyer & A. Wood, Trans.). Cambridge: Cambridge University Press.
- Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin*, 122(3), 250–270.
- Kipling, R. (1912). *Just So Stories*. New York: Doublday Page & Company.
- Kirkpatrick, L. A. (1999). Toward an Evolutionary Psychology of Religion and Personality. *Journal of Personality*, 67(6), 921–952. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=keh&AN=2697423&site=ehost-live>
- Kitayama, S., Varnum, M. E. W., & Salvador, C. E. (2019). Cultural Neuroscience. In D. Cohen & S. Kitayama (Eds.), *Handbook of Cultural Psychology, Second Edition*. New York: Guilford Press.
- Kitcher, P. (2006). Ethics and evolution: How to get here from there. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 120–139). Princeton, NJ: Princeton Univeristy Press.
- Kohlberg, L. (1963). The development of children's orientations toward a moral order: I. Sequence in the development of moral thought. *Vita Humana*, 6(1–2), 11–33.
- Koltko-Rivera, M. E. (2006). Rediscovering the later version of Maslow's hierarchy of needs: Self-transcendence and opportunities for theory, research, and unification. *Review of General Psychology*, 10(4), 302–317. doi:10.1037/1089-2680.10.4.302

- Korsgaard, C. M. (2006). Morality and the distinctiveness of human action. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 98–119). Princeton, NJ: Princeton University Press.
- Korsgaard, C. M. (2010). *Reflections on the Evolution of Morality*. Amherst Lecture in Philosophy. Retrieved from www.amherstlecture.org/korsgaard2010/
- Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science*, 3(3), 149.
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819–834.
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58, 317–344. doi:10.1146/annurev.psych.58.110405.085658
- Lewis, G. J., & Bates, T. C. (2010). Genetic Evidence for Multiple Biological Mechanisms Underlying In-Group Favoritism. *Psychological Science*, 21(11), 1623–1628. doi:10.1177/0956797610387439
- Lord, R. G., Diefendorff, J. M., Schmidt, A. M. et al. (2009). Self-regulation at work. *Annual Review of Psychology*, 61, 543–568. doi:10.1146/annurev.psych.093008.100314
- Loyau, A., Petrie, M., Saint Jalme, M. et al. (2008). Do peahens not prefer peacocks with more elaborate trains? *Animal Behavior*, 76, e5–e9. doi:10.1016/j.anbehav.2008.07.021
- Maslow, A. H. (1969). The farther reaches of human nature. *The Journal of Transpersonal Psychology*, 1(1), 1–9.
- McAdams, D. P. (2009). The moral personality. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 11–29). New York: Cambridge University Press.
- McAuliffe, K., Bogese, M., Chang, L. W., et al. (2019). Do dogs prefer helpers in an infant-based social evaluation task? *Frontiers in Psychology*, 10, 591. doi:10.3389/fpsyg.2019.00591
- McElreath, R. (2010). The coevolution of genes, innovation and culture in human evolution. In P. M. Kappeler & J. B. Silk (Eds.), *Mind the Gap: Tracing the Origins of Human Universals* (pp. 451–474). Berlin: Springer.
- McKinnon, S. (2005). *Neo-Liberal Genetics: The Myths and Moral Tales of Evolutionary Psychology*. Chicago: Prickly Paradigm Press.
- McRae, K., Ochsner, K. N., & Gross, J. J. (2011). The reason in passion: A social cognitive neuroscience approach to emotion regulation. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of Self-Regulation: Research Theory, and Applications* (pp. 186–203). New York: Guilford Press.
- Mead, G. H., & Morris, C. W. (1934). *Mind, Self, and Society from the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, 11(4), 143–152.
- Miller, G. F. (2007). Sexual selection for moral virtues. *The Quarterly Review of Biology*, 82(2), 97–125.
- Miller, H. C., Pattison, K. F., DeWall, C. N. et al. (2010). Self-control without a “self”?: Common self-control processes in humans and dogs. *Psychological Science*, 1–5.
- Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: a working model. *Neuroreport*, 14(3), 299–305. doi:10.1097/OL.wnr.0000057866.05120.28
- Moll, J., De Oliveira-Souza, R., & Zahn, R. (2009). Neuroscience and morality: Moral judgments, sentiments, and values. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 106–135). New York: Cambridge University Press.
- Oatley, K., Keltner, D., & Jenkins, J. M. (2006). *Understanding Emotions*. New York: Wiley-Blackwell.

- Ochsner, K. N., Ray, R. R., Hughes, B. L. et al. (2010). Bottom-up and top-down processes in emotion generation. *Psychological Science*, 20(11), 1322–1331.
- Peterson, C., & Park, N. (2010). “What happened to self-actualization” commentary on Kenrick et al. (2010). *Perspectives on Psychological Science*, 5(3), 320–322. doi:10.1177/1745691610369471
- Pietraszewski, D., & Wertz, A. E. (2021). Why evolutionary psychology should abandon modularity. *Perspectives on Psychological Science*, 17(2), 465–490. doi:10.1177/1745691621997113
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1), 1–20. doi:10.1017/s0140525x02000018
- Prinz, J. J. (2008). Is morality innate? In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (vol. 1, pp. 367–406). Cambridge, MA: MIT Press.
- Profet, M. (1988). The evolution of pregnancy sickness as protection to the embryo against Pleistocene teratogens. *Evolutionary Theory*, 8, 177–190.
- Rawls, J. (1971/1999). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Richerson, P. J., & Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: University of Chicago Press.
- Richerson, P. J., Boyd, R., & Henrich, J. (2010). Colloquium paper: Gene–culture coevolution in the age of genomics. *Proceedings of the National Academy of Sciences*, 107(Supplement_2), 8985–8992. doi:10.1073/pnas.0914631107
- Robins, R. W. (2021). Naturalizing the Self. In O. P. John & R. W. Robins (Eds.), *Handbook of Personality: Theory and Research* (4th ed., pp. 561–586). New York: Guilford Press.
- Rodrigues, S. M., Saslow, L. R., Garcia, N. et al. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences*, 106(50), 21437–21441. doi:10.1073/pnas.0909579106
- Schaller, M., Neuberg, S. L., Griskevicius, V. et al. (2010). Pyramid Power: A Reply to Commentaries. *Perspectives on Psychological Science*, 5(3), 335–337. doi:10.1177/1745691610369474
- Schmitt, D. P., & Pilcher, J. J. (2004). Evaluating evidence of psychological adaptation: How do we know one when we see one? *Psychological Science*, 15(10), 643–649.
- Schwartz, S. H. (2006). Basic human values: Theory, measurement and applications. *Revue française de sociologie*, 4(47), 929–968.
- Shweder, R. A. (2012). Relativism and universalism. In D. Fassin (Ed.), *A Companion to Moral Anthropology* (pp. 85–102). New York: Wiley-Blackwell.
- Shweder, R. A., Mahapatra, M., & Miller, J. G. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The Emergence of Morality in Young Children* (pp. 1–83). Chicago, IL: University of Chicago Press.
- Skitka, L. J., & Mullen, E. (2002). The dark side of moral conviction. *Analyses of Social Issues and Public Policy*, 2(1), 35–41.
- Sober, E. (2006). *Conceptual Issues in Evolutionary Biology* (3rd ed.). Cambridge, MA: MIT Press.
- Sober, E., & Wilson, D. S. (1999). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Spencer, H. (1866). *The Principles of Biology* (vol. 1). New York: Appleton & Company.
- Stich, S. P. (2006). Is morality an elegant machine or a kludge. *Journal of Cognition and Culture*, 6(1–2), 181–189.
- Thompson, N. S. (1998). Reintroducing “reintroducing group selection to the human behavioral sciences” to BBS readers. *Behavioral and Brain Sciences*, 12, 304–305.
- Tse, P. U. (2008). Symbolic thought and the evolution of human morality. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Evolution of Morality: Adaptations and Innateness* (vol. 1, pp. 269–297). Cambridge, MA: MIT Press.
- Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. In M. A. S. Killen, Judith (Ed.), *Handbook of Moral Development* (pp. 7–36). Mahwah, NJ: Lawrence Erlbaum Associates.

- Van de Vondervoort, J. W., & Hamlin, J. K. (2018). The early emergence of sociomoral evaluation: infants prefer prosocial others. *Current Opinion in Psychology*, 20, 77–81. doi:10.1016/j.copsyc.2017.08.014
- Walker, L. J., & Frimer, J. A. (2015). Developmental trajectories of agency and communion in moral motivation. *Merrill-Palmer Quarterly*, 61(3), 412–439.
- Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who sees human?: The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3), 219–232. doi:10.1177/1745691610369336
- Waytz, A., Gray, K., Epley, N. et al. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, 14(8), 383–388. doi:10.1016/j.tics.2010.05.006
- Weinstein, D. (2009). Herbert Spencer. *Stanford Encyclopedia of Philosophy* (Summer 2009). Retrieved from <http://plato.stanford.edu/archives/sum2009/entries/spencer/>
- Weisfeld, G. E., Czilli, T., Phillips, K. A. et al. (2003). Possible olfaction-based mechanisms in human kin recognition and inbreeding avoidance. *Journal of Experimental Child Psychology*, 85(3), 279–295. doi:10.1016/s0022-0965(03)00061-4
- Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, 15, 728–755. doi:10.1080/17405629.2018.1435413
- Whiten, A., & van Schaik, C. P. (2007). The evolution of animal “cultures” and social intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 603–620. doi:10.1098/rstb.2006.1998
- Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, 17(4), 585–654. doi:10.1017/s0140525x00036104
- Wilson, D. S., & Sober, E. (1998). Multilevel selection and the return of group-level functionalism. *Behavioral and Brain Sciences*, 21(2), 305–306. doi:10.1017/s0140525x98221194
- Wright, J. C. (2021). Morality as a regulator of divergence: Protecting against deviance while promoting diversity. *Social Cognition*, 1, 81–98.

Neuroscience of Moral Action

Henry Gustav Molaison, though never a scientist, has contributed to psychological neuroscience more than most neuroscience researchers. He did so in his role as H. M., the celebrated epilepsy patient who had large sections of his hippocampus and amygdala removed, along with portions of his cortex near these structures. After the surgery, his memory was dramatically altered: in addition to acquiring amnesia, he seemed to have no ability to remember new things. Over the next fifty-five years of his life, he was unfailingly hospitable and welcoming to the troops of researchers who repeatedly visited him, although he could never really remember any of them.

H. M.'s contributions to neuroscience are usually categorized in the area of cognition and learning, but his peculiar memory deficit was a profound social hindrance. Every social experience was a new one, even those with researchers who had known him most of his life. He could not hold a job and lived "entirely in the moment" (Carey, 2008). This social aspect of H. M.'s disadvantage reminds us that, for all our vaunted cognitive ability, humans are a profoundly social species. Much of our social, and thus in large part moral, behavior is embedded in the underlying neurological and biological processes that we share with other mammals (de Waal, 2006).¹ In addition to this shared heritage, humans add the complexities of reason, self-consciousness, and social consciousness, as well as our endeavors in science, philosophy, religion, etc.

It is to these complexities that we now turn. First, we provide an overview of how one might understand these complexities to be embedded and embodied or structured and supported by neurological underpinning. How are neuroscientists conceiving this relationship? Based on those conclusions, we then provide an overview of the neuroscience of moral action that is structured by the functions the neural systems must serve (rather than by the neuroanatomy). We take the functional categories of moral action from the chapter structure of this text, then inquire into how the neural systems underlie, support, or embody those influences and processes in the chapters.

¹See Chapter 1 for an exploration of the crucial differences in morality in humans and other mammals.

2.1 Neural Systems and Moral Function, or How to use Neuroscience

This chapter is an attempt to draw some lessons from neuroscience about how human morality functions. If we want to use neuroscience to understand moral psychology, we need two things: (1) good psychological theory and data about aspects of the processes themselves; and (2) good neuroscientific models and data about the locations and interconnections of the systems that support them (Kihlstrom, 2007). These two sources can work interactively to help us construct an increasingly better picture of the biological, psychological, social, and cultural processes that constitute our morality. In this collaboration, it is the psychological theory that describes what we (think we) know about the function, and the neuroscience that maps the location and interconnections of the neural systems supporting that function in the brain.² In turn, understanding which locations and interconnections (that is, which systems) are active in supporting that function, over the time process of the function, can help us make better inferences about that function.³

For instance, one *might* ask what moral functions the amygdala serves; *we ask rather* what neural systems are involved in showing empathy. Thus, we begin with the psychological function (e.g. empathy) and move to the neuroscience (e.g. a shifting system of locations that often but not always includes the amygdala, over the varied time processes and varied experiences,⁴ and how those systems work in synchrony with systems in other people's brains). This frees us from thinking that any single location (or even regular set of locations) is "the place" where that function occurs. We still cite work that implicates particular locations since those offer clues to the system(s), their interconnections, and the unfolding of the process over time.

For instance, Lebois et al. (2020) have shown that different neural systems with only mild overlap (as little as 11.5%) are responsible for social-exclusion fear vs. physical-harm fear and social-exclusion anger vs. physical-harm anger. The psychological theory (see Chapter 8) is that these are situated forms of anger and fear, and

²Kihlstrom (2007) gives a good example of this with the saga of the famous H. M. As psychological theory progressed, the explanation for H. M.'s symptoms changed over several decades: from (1) learning capacity loss; to (2) long- but not short-term memory loss; to (3) deep but not shallow processing loss; to (4) declarative but not procedural memory loss; then most recently to (5) relational but not non-relational memory loss. Almost all this change in the description of the function of the Hippocampus and amygdala was driven by psychological and not neuroscientific research. Thus, the attributed functions of the areas changed, not because of advances or new findings in neuroscience, but because of our changing understanding of the functions.

³Most reviews of the neural substrates of morality are organized based on the neural structures themselves (Casebeer, 2003; Greene & Haidt, 2002; Mendez, 2009; Moll, de Oliveira-Souza, & Eslinger, 2003; Moll et al., 2009). One can find a few reviews of moral neuroscience organized, like ours, on function (Barrett, 2017; de Oliveira-Souza, Zahn, & Moll, 2016; Narvaez, 2008; Narvaez & Vaydich, 2008).

⁴Remembering that the experience we designate with a single word (empathy, anger, fear) is really a family of experiences – there is *fear* of physical danger and *fear* of social negative evaluation (Lebois et al., 2020) and they look quite different in their neuroscientific embodiment.

that the differences in the neurological underpinning are driven by the *meaning* of the situation that drives the type of fear or anger. So here, it seems that meaning is shaping or structuring (or even constructing) the neural system that underpins the specific emotion. This is a move away from simple mental categories and toward how they are embedded in particular situations and complex behaviors (Pessoa, 2022; Pessoa, Medina, & Desfilis, 2022).

Thus, we need to take a moment for a note on philosophy of science before beginning our survey. Too often those who review the neuroscience of morality (or of anything, really) see neuroscience as a talisman of truth. We are all taught that correlation does not imply causation, and then promptly forget it when looking at hot spots on an fMRI (functional magnetic resonance imaging) brain scan.⁵ We attempt to be careful with causal language in this chapter (and indeed in all chapters). Neuroscience is amazingly useful in both science and technology, but we too often (1) see the neural structures and processes as the “real” structures and processes; and (2) see the psychology (and anthropology, and philosophy, and religion) as things that can be reduced to the neuroscience. These two mistakes have been termed neuro-realism and neuro-essentialism (Racine et al., 2005). This involves making philosophical assumptions (e.g. only one level of explanation is “real”) that are often unspoken and undefended. The current state of the field does not really support any simple version of reduction of mind to brain. Thus, rather than saying the physiology causes the psychological function, we prefer to speak of it underlying, supporting, or embodying the function.⁶

One final and related point: any search for a “moral center” that is said to “control” some function of the brain disregards the massive interactivity of all the various brain structures (Pessoa, 2008, 2009; Sinnott-Armstrong, 2016) and the likely distribution of any function across many parts of the brain. Consider the career of the hormone/neurotransmitter oxytocin. Once talked of as the “love hormone” its effects have become clouded by reports that it has contradictory effects, sometimes affiliative and sometimes antisocial (Alvarez, Quintana, & Whitehouse, 2016). This puzzle is beginning to be resolved by the work that shows that psychological systems condition its effects. For instance, individuals with insecure attachments in childhood can respond to oxytocin increases with antisocial behavior, perhaps because of the fear that affiliation engenders in them (Bartz, 2016; Bartz et al., 2015). Mierop et al. (2020, p. 1228) in reviewing the field, suggest that it is still “virtually impossible” to know which of these effects are replicable at this point.

So, now we turn to a review of various aspects of, contexts for, influences on, and processes of moral action as represented by the chapters of this text.

⁵The methodological issues associated with analysis and interpretation of fMRI data are legion (for examples of these difficulties, see Poldrack & Mumford, 2009; Racine, Bar-Ilan, & Illes, 2005; Vul et al., 2009). Postle (2016) suggests that increased activity in an area may not be diagnostic of the use of information associated with that area (e.g. emotional or relational information) and Elliot et al. (2020) provide meta-analytic evidence that test-retest reliabilities are far too low to support the prediction of clinical outcomes.

⁶See Brigandt and Love (2017) for a further philosophical explanation of reductionism and some of its alternatives and (Barrett, 2017) for a complex version of determinism that is not reductionist.

2.2 Moral Ecology

In an article on the emerging field of cultural neuroscience, Ambady and Bharucha (2009) called the brain “the organ of culture.” The brain is indeed at least one place where the patterns and influences of our social surround encounter the patterns and proclivities of our inner life. To be embodied in the individuals who make up a moral ecology, the social complexities we describe in Chapter 3 still need to be channeled through those people. This channeling is done by processes of imitation, social influence, and social cognition, and is supported by a network of interacting brain processes (and interacting brains) that neuroscientists are only beginning to map. In this section, we will review a subset of what we know about how moral ecology influences individuals through imitation, social cognition, and social influence.

2.2.1 Imitation and Mirror Neurons

Humans have a ubiquitous tendency to imitate one another. We align our behavior with others in both literal (e.g. posture, speech) and symbolic (e.g. stereotype-based) ways, and do so unintentionally, even without conscious awareness (Iacoboni, 2009). This is a very basic and pervasive form of social influence. It is supported by a complex network of action perception and interpretation that involve one of the most fashionable topics in neuroscience: mirror neurons. These are neurons that are part of the sequence involved in controlling motor movement for the macaque monkeys in the original study (moving the hand to the mouth to eat) but that were also found to activate when the monkey watched the experimenter perform the same action (di Pellegrino et al., 1992). Thus, they “mirrored” the perceived activity in a section of the brain devoted to physically performing the activity. Accumulated research in humans since this initial finding now shows that there is a section of the superior temporal sulcus that processes intentional movement in others and that connects to two separate areas of mirror neurons, one that represents the motor aspects of the movement (parietal mirror neuron area) and one that represents the goal aspects of the movement (frontal lobe mirror neurons) (Iacoboni, 2009). It seems clear at this point that mirror neuron brain areas contribute to imitative action (e.g. copying body movement) but not directly to higher level processes like inferring intentions (Campbell & Cunningham, 2017; Grigaityte & Iacoboni, 2016; Heyes & Catmur, 2022). Tracking how mirror neurons are integrated into larger systems is work that remains to be done.

2.2.2 Thinking About, and With, Others

One of the central psychological processes that underlies social influence is our ability to think about others, to take their perspective, and to recognize that they are judging us. This ability requires representing the mental states of others. Again, we find that the medial prefrontal cortex (mPFC) is involved in this, the same structure that is centrally involved in processing information about the self (Heatherton, 2011). There seems to be specialization within the mPFC, with other-mind processing areas toward the dorsal (top) side of the mPFC and self-processing areas toward the ventral (bottom) side. Interestingly, there are some clues that the processing may differ for

in-group and out-group members, with activation reported in both the self-processing and other-processing areas for people who are considered similar to the self (Heatherton, 2011).

Others also influence us because we are concerned about judgments they might make about us. This awareness of other people can be threatening and emotional, and thus – in addition to mPFC – the awareness of threat from others involves areas that modulate and tune emotional expression. We have covered this social threat detection in Section 2.4 and processes of empathy in Section 2.8.

In a significant advance in theory and method, Tamir and Thornton (2018) have paired advanced techniques in neuroscience imaging with a dimensional model of social perception that looks at three levels of social perception:⁷ perception of character traits of the other, perception of the current state of the other, and evaluations of the actions of the other. This is a complex and promising model of how social inference plays out at the neural level. It is another example of the massive interactivity evident at both the psychological and neural levels.

2.2.3 Cultural, Group, and Individual Influence on Brain Processes

We have already seen that psychological characteristics can modify the way the brain processes information.⁸ The mPFC in individuals with low self-esteem processes information about others differently and information about out-groups is processed differently than information about in-groups (Heatherton, 2011; Kitayama, Varnum, & Salvador, 2019; Somerville, Heatherton, & Kelley, 2006). The extent to which a culture is tight (*vs.* loose⁹) is associated with more reactivity in brain areas linked with norm violation detection (Kitayama et al., 2019). There also seems to be evidence that the neuroanatomy associated with thinking about the self is different in different cultures (Kitayama et al., 2019). Zhu et al. (2007) have shown differences in Chinese and Western college students' neural representations of self and mother.¹⁰ As mentioned before, mPFC is associated in Western studies with active representation of the self, but in this study Chinese students' active representations of their mothers also showed significant activity in mPFC (among other cultural differences in processing). Thus, the functions associated with particular areas of the cortex (the functional anatomy) differ in these samples, based on culture. For Chinese people, there is “neural unification” of the self and the mother, while for Westerners there is “neural separation” (Zhu et al., 2007, pp. 1314–1315). Western culture is, on average, much more

⁷They use representational similarity analysis in fMRI. This advanced analytic approach allows comparison of the similarity of reaction in entire neural systems across people who have been shown similar stimuli. With carefully chosen stimuli, this allows one to see similarities in neural response patterns that represent the psychological dimension (e.g. perceptions of social power).

⁸This should only be surprising to those who see neural events as the single “real” level of explanation (see neuro-essentialism in Section 2.1).

⁹The extent to which a culture requires close observance of morals and customs or allows a looser relationship to them. See Chapter 9, Section 9.3.1.3.

¹⁰These studies are done by looking for brain areas that are activated when people are asked to think about themselves, about their mother, about important others, etc.

individualistic than Chinese culture, which tends to see others as forming a very close bond with the self. Thus, the “neural unification” of an important other suggests a cultural influence on functional neuroanatomy. This effect is complicated when one looks at Chinese people with significant exposure to Western, individualistic culture (Kitayama et al., 2019)

2.2.4 Interpersonal Neuroscience

It is now possible to track how two or more brains show synchrony in patterns of activity, opening the possibility of a truly interpersonal neuroscience (Pan, Novembre, & Olsson, 2021).¹¹ This brain-to-brain synchrony (BB synchrony) has been shown to strongly correlate with social learning in dyads and even at the classroom level (Pan et al., 2021). Facial, voice, and movement cues seem foundational in this process of establishing synchrony, and synchrony occurs at behavioral, affective, perceptual, and neural levels (e.g. emotional contagion is associated with neural synchrony or entrainment) (Wheatley et al., 2011). BB synchrony, but not self-reported group identification, has been shown to predict team performance in both an economic game and problem-solving tasks. Basso, Satyal, and Rugh (2020) collate evidence suggesting that BB synchrony is important in shared dance, helping to explain the social synchronizing function of shared dance and movement (e.g. ritual, folk dance, etc.) across cultures.

2.2.5 The Mutual Constitution of Brain and Cultural Processes

Kitayama et al. (2019, p. 80), in a review of cultural neuroscience, collate the empirical evidence for a claim long dominant in cultural anthropology. Their summary of the claim is worth quoting at some length:

the body (and now the brain as well) is *closely attuned to the sociocultural environment*, while at the same time ... autonomous, *thereby constituting agency*, being capable of producing volitional actions, which can lead to changes in the environment from which the agency has been derived. This circular or recursive process occurs continuously, not only in each person's lifetime but also across generations, ... giving rise to changes in both historical and evolutionary timescales. [emphasis ours]

The central point here is the close attunement and mutual influence of brain–body systems and cultural systems. Cultural systems both resist change and provide resources for individuals and groups to produce change. Individuals can both resist and seek individual change using resources available to them from culture. Each one influences and structures the other over both short and long timescales.

¹¹The methods for tracking synchrony involve some measurement of activity, e.g. electroencephalogram (EEG) or fMRI or functional near-infrared spectroscopy (fNIRS, a less invasive approach in which the subject wears a cap that tracks brain activity in a similar, but less detailed, way to fMRI). Mathematical models are then used to highlight parallelism in activity between two or more brains in people who are interacting.

In summary, we can see that social influence and moral ecology are attuned to brain processes and systems at many different levels. This ranges from low-level mirroring of motor action, to the mirroring of feeling states, to fears of social sanction from others and cognition about others to synchronization of movement, affect, and intention. This attunement to moral ecology provides the support and channel for effective volitional action within that moral ecology.

2.3 Personality

The attempt to connect personality to physiological processes is as old as the classical Greek theories of the humors (Zuckerman, 1995). The human body was thought to contain four kinds of humors (fluids) and various personality traits (e.g. melancholy) were thought to be associated with imbalances in these humors. But it is only with the recent, rapid development of neuroscience measurement techniques that we have been able to make some empirical progress in actually evaluating proposed theoretical connections.

We call the enduring patterns of how individuals relate to the world their personality. This corresponds to the organizing question in Chapter 4: What does a person like me do in a situation like this?¹² These enduring patterns structure moral action at three levels. *Traits* (e.g. extraversion, cynicism) describe broad patterns of individual consistency across time and situation. *Characteristic adaptations* (such as personal projects, life goals) provide a more fine-grained analysis of patterns of consistency. A final level of personality consists of the patterns in the *narratives* we construct to seek out and identify meaning in our lives. These are unique to each individual, but one can find patterns even here – e.g. people differ in the kinds of themes they include in their life stories (McAdams, 2009).

Most of the work we review has been done at the level of traits, and almost nothing has been done at other levels of personality, such as narrative. De Young and Gray (2009) provide a useful overview of various theorists' claims about the relation of personality to neuroscience and state that it is simply too early in personality neuroscience to link neural structure to the more specific levels of personality – e.g. personal adaptations and narrative (McAdams & Olson, 2010). These levels are heavily dependent on personal developmental trajectory and unique cultural surround and it will likely be difficult, though not impossible, to connect them to neural structure. Markett, Montag, and Reuter (2018) propose a move in personality neuroscience similar to that documented in Section 2.2: looking for the patterns not in particular brain structures but in networks of brain structures as they interact across time. In a movement in that direction, Wright et al. (2019) provide evidence that differences in serotonin functioning (and thus the systems implicated in this functioning) underlie differences on the personality meta-trait of stability.¹³ But beyond this, we are currently limited to looking at associations with particular structures.

¹²Both how I characteristically view myself and how I characteristically view situations are enduring aspects of my personality.

¹³Consisting of three – conscientiousness, agreeableness, and neuroticism (reverse coded) – of the “Big-5” personality traits.

Although alternative models are still occasionally floated (Zuckerman, 1995), the most widely accepted structure for a trait-level description of personality is the “Big 5” (McCrae & Costa, 1997). Recent work with structural MRI scans by deYoung et al. (2010) provides a comprehensive view of the relationship of the size¹⁴ of particular neural structures associated with differences in the Big 5 personality traits, replicating findings from other work. They found clear associations for four of the Big 5 traits (openness to experience had less firm, but still suggestive, patterns). For instance, individuals high in *extraversion* tended to have a larger medial orbitofrontal cortex, an area involved in assessing the reward value of stimuli. Those individuals scoring higher in *agreeableness* tended to have smaller posterior left superior temporal sulci (pSTS) – an area associated with inferring intention based on patterns of movement – and to have a larger posterior cingulate (PC) – an area associated with understanding other’s beliefs. The functions associated with these areas are crucial perceptual and cognitive components of empathy. Individuals high in *conscientiousness* tended to have larger middle frontal gyri – an area of the frontal cortex that is central in maintaining working memory and executing planned action. Finally, individuals high in *neuroticism* had a more complex pattern of structure, with some structures being smaller and others larger but all associated with emotional sensitivity to threat and punishment (DeYoung et al., 2010).

Note the pattern of the scientific inference in all these claims: (1) a correlation of a personality trait with size of a structure is reported (agreeableness with the pSTS); (2) the structure itself is then characterized in terms of other functions that might explain the correlation (tracking intention is associated with pSTS). Thus, we know that higher scores on both agreeableness and altruism are related to a brain structure associated with inferring intention (Tankersley, Stowe, & Huettel, 2007). But that does *not* make the pSTS the “altruism center.” Tracking intention and understanding other’s beliefs are only a part of the larger complex of coordinated activity that supports altruistic action. The larger personality aspect of altruism is not identical to or reducible to the specific brain structure. But we can tie particular aspects of the neurological activity to specific aspects of the psychological function.

More recent work that maps psychological constructs at the neural level shows how extraordinary altruism is supported by the functional neural processes that underlie empathy. Individuals who have taken the unusual step of donating a kidney to a stranger have been shown to have a greater overlap of activation in the anterior insula for fearful stimuli not only for the self *but also for others* (as compared to matched controls). The suggestion is that this activation is one aspect of an “affective salience network” that is more responsive to harm in others than usual (Brethel-Haurwitz et al., 2018, p. 1638). Again, we see aspects of psychological processes for empathy underpinned by distributed networks associated with psychological functions.

¹⁴DeYoung (2010) provides a useful discussion about whether the volume of an area is the right metric for measuring its relative performance of a function. We agree that simple interpretations of volume leave much to be desired, and remind one of earlier, discredited phrenological approaches. But we also note that there does seem to be a replicable association of size with personality functions. As with all work in this area, scientific humility is in order.

2.4 Moral Identity and the Self

The moral self-concept, or moral identity, has begun to play an increasing role in explanations of why people take moral action. There are two aspects of the self that allow it to become involved in influencing moral action: (1) the content and structure of self-knowledge; and (2) those self-conscious emotions (pride, embarrassment) that guide action and action planning.¹⁵ Self-knowledge and self-conscious emotion are linked because we can think about ourselves over time and reflect on our reactions to evaluations by ourself and by others (Leary, 2007).

One particular brain structure is closely associated with tasks that involve moral self-concept: the medial prefrontal cortex (mPFC). Still, though self-related tasks show heightened mPFC activity, there are few aspects of the moral experience that can be isolated to one specific region of the brain. The mPFC by itself cannot be called the location of the self. In a review paper section titled “Is the mPFC the self?” Heatherton (2011) explains: “the experience of the self involves various sensory, affective, and motor processes contributed by disparate brain regions outside the cortical midline area [containing the mPFC]” (p. 370). For example, an early finding identified the mPFC as the most active when Western individuals were asked to think about themselves in terms of abstract personality traits. But when this same task was given to a Chinese sample, the activation in the mPFC was much weaker. This is consistent with the idea that individuals with an interdependent conception of the self (one tightly connected to others) do not think of themselves as much in terms of abstract personality traits (Kitayama et al., 2019).¹⁶

In short, to find those aspects of the brain that underlie the experience of the self, one must look to a system of structures, processes, and functions, some of which may be distributed across multiple individuals (via e.g. BB synchrony) and cannot be precisely localized even in any one brain.

2.4.1 Self-Knowledge

Research has shown that the mPFC is strongly related to semantic knowledge about the self. Most studies that compare thinking about the self with thinking about other things or people identify the mPFC as showing heightened activity for thinking about the self.¹⁷ Beer (1999) proposes that the high “resting metabolic rate” of the mPFC suggests that it is involved in the constant self-evaluation that accompanies our

¹⁵Non-self-conscious or outwardly focused emotions (e.g. vicarious anger, empathy) can also guide moral action, and their ability to promote moral action may be related to self-concept in a more complex manner (e.g. by reducing inhibition to action). However, these influences are speculative, and there is little to no neuroscience research that investigates them.

¹⁶We should note that though this assumes collectivist societies consist of many people with interdependent views of the self, the picture is more complicated. Latin American societies are strongly collectivist, but they also emphasize an independent view of the self (Krys, 2022)

¹⁷See Heatherton (2011) for a review and Mitchell et al. (2005) for divergent findings that show mPFC occasionally involved in judging others, even dogs.

everyday activity. The range of self-evaluative tasks that have been shown to involve the mPFC (free-form reflection, self-referential memory, self-judgment all show activation of mPFC; mindfulness meditation shows deactivation) and the deficits shown with damage to this area (memory for personal preferences and other self-knowledge) all suggest that this area is central in processing knowledge (and likely moral knowledge) about the self (Heatherton, 2011). One must be careful in making broad claims. Kitayama et al. (2019) review evidence that these patterns differ markedly across cultures, and that the ventro-mPFC (vmPFC) is less active when thinking about the self in cultures where the self is typically interdependent, though in Latin American cultures these aspects seem to be independent (Krys, 2022).

2.4.2 Self-Conscious Emotion

Self-conscious emotion is a response to an appraisal of some threat to the self or of some incentive for the self (see Chapter 8 for how appraisal works). This distinction between threat and incentive is based on Higgins's (1997) distinction between promotion and prevention goals. Promotion goals seem to be associated with heightened activity in the left orbital and mPFC, areas associated with processing reward information and self-referential knowledge respectively (Eddington et al., 2007; Strauman et al., 2012).

Not surprisingly, the amygdala is involved when threats are detected and prevention becomes important. But it seems to be the anterior cingulate cortex (ACC) – immediately adjacent to the mPFC with its self-relevant functions – that regulates the connection between the mPFC and the amygdala in the regulation of threatening social emotion (Heatherton, 2011; Strauman et al., 2012). Thus, the psychological processes associated with promotion and prevention goals seem to be subserved by *different* neural circuitry and therefore the evaluation and pursuit of these different goals may differ qualitatively.¹⁸ But work needs to be done before one can make such strong conclusions.

To complicate matters, these brain areas do their work in concert with a range of other areas. For instance, another area in the frontal cortex (the orbito-frontal cortex, OFC) seems to be involved in the appraisal process for social emotions (Beer, 2007; Beer, Lombardo, & Bhanji, 2010). Damage to this area produces patients who can generate social emotions but often the inappropriate ones. Yet another complication is that the ACC seems to do its threat-detection work differentially in those with low self-esteem (much more active in response) than in those with normal or high self-esteem (Somerville et al., 2006).

Thus, it seems we have a complex of areas in the midline of the cortex (mPFC, ACC, OFC, but also amygdala) that are centrally involved in mediating the self's involvement in moral perception, reaction, and action. They act in concert but they are not isolated from much other activity. We are only beginning to understand how this activity might be related to the self and to moral action. Constructing that understanding will likely involve identifying shifting neural systems over the timescale of the experience.

¹⁸See Chapter 9 for more explanation about these different motivational and self-regulatory systems.

2.5 Skills and Knowledge

Since moral action implies agency, skill is inevitably implicated. Skill is involved not only in executing action but also in perceiving the need or opportunity for it, evaluating possibilities, and planning for action. Skill is also involved in the emotional regulation associated with effective social action, and its neural underpinning can differ depending on the shape emotional regulation takes in different cultures (Kitayama et al., 2019, p. 95).

It is well established that increases in training (e.g. comparing novices to experts in some area) can produce increases in the size, density, and/or reactivity of related brain areas that serve the skill (one example of a general process called neuroplasticity). This has been replicated in music (Herdener et al., 2010; Kleber et al., 2009), motor skill and motivational aspects of skill training (Boyke et al., 2008), perception (Herdener et al., 2010), and even compassion meditation (Lutz et al., 2009; for a review, see Tang, Holzel, & Posner, 2015).

As a habit develops based on repeated practice, its neural underpinnings shift from the cortical networks associated with goal-driven action control to the habit-based sensorimotor cortico-basal ganglia (Wood, Mazar, & Neal, 2021, p. 3). Computational models suggest that movement back and forth between goal-based and habit-based action could take multiple forms, allowing one to think of the two systems as either separate or as tightly conjoined (Wood et al., 2021, p. 4). Basso et al. (2020) review evidence that dance and associated BB synchrony facilitate group cohesion and that the cortico-spinal tract in dancers and athletes shows increased organization. This suggests yet another expansion of neural areas related to moral action.

The evidence that different kinds of expertise produce a wide variety of changes at the neuroanatomical level suggests that we need to expand dramatically our map of those brain areas that support the wide variety of moral action. To the extent that particular skills and expertise are inextricably intertwined with moral goals (e.g. compassion, the design of safe computing systems, supporting group functioning) it is difficult to argue that the neural systems that support these skills are not parts of the “moral brain.” This expansion of the territory of the moral brain opens the door to a host of other neuroscience findings in cognition, learning, habit, expertise, and emotion. As moral judgment and decision-making is pushed into automatic skills and habits, so also their neural underpinnings must be expanded into those brain areas serving those processes. This expansion may be so large that it might be better to stop talking about particular areas being most associated with abstract-level morality.¹⁹

2.6 Moral Reason

Long considered the paragon and defining characteristic of morality, reason has been somewhat under siege (Haidt, 2010; Narvaez, 2010). The recent spate of neuroscience findings linking emotion centers to moral judgment has been one of the driving

¹⁹For instance, Ellemers and van Nunspeet (2020) review work showing that conformity changes brain activity associated with visual perception, indicating that social influence may change the way we perceive the world, and providing yet another expansion of the moral brain.

forces in this assault (Greene & Haidt, 2002; Greene et al., 2001). Here we review work that tries to isolate or characterize the role of reason in moral action by identifying its neural correlates.

2.6.1 Pure Reason

Work in the neuro-economics tradition by Shenhav and Greene (2010) tracks the neural underpinnings of the mental calculation of expected moral value, a combination of the *magnitude* of a possible moral loss/gain and the *probability* of that loss/gain. Changes in expected moral value are particularly related to the vmPFC, and the medial OFC (mOFC). In this study, individuals made judgments about expected moral value while being presented with a range of possible values of magnitude in gain/loss and probability. This design allowed the researchers to show that these structures were particularly sensitive to changes in these values. The pattern of results suggests that these calculations of moral value are done using the same processes that other work has found to be active in more straightforward self-interested economic decision-making. This particular value tracking is, they argue, not a specialized *moral* calculation but a more general-purpose process.

Shenhav and Greene (2010) also tracked the individual differences that might lead one person to be more utilitarian in their approach than another. Individuals who are more likely to hew closely to utilitarian judgments show more activation in the OFC. This activation seems to be in line with the association of the OFC in identifying and modulating emotional reactions (Beer, Knight, & D'Esposito, 2006). Perhaps the people doing pure utilitarian calculations were better at blocking out emotional reactions.

2.6.2 Balancing Reason and (Social/Emotional) Reward

The utilitarians in Shanhav and Greene's (2010) study seem to be blocking out emotional input in favor of a stark, mathematical moral calculus. But clearly others *are* including emotional aspects in their judgment. Is including emotion in moral judgment a good thing, or should all moral judgment instead resemble the pure utilitarians just discussed?

Extensive work in developmental psychology on risk-taking in adolescence can provide some enlightenment here. Why are adolescents more willing to take risks, particularly when they are together with their friends? There appear to be two different systems involved in calculations of risk: a *socio-emotional* dopaminergic reward-seeking system (centering on the limbic system) and a *cognitive control* network involving planning, thinking ahead, and self-regulation (Steinberg, 2007). The emotional system develops earlier, and excessive risk-taking in adolescence seems to be associated with the *imbalance* between socio-emotional reward for risk and cognitive control that might (but cannot yet) modulate it. It is achieving the *balance* between these systems that produces adult-level risk-taking. Adolescent reasoning ability, considered alone, is comparable to adults, but adolescents find the psychosocial rewards associated with risk (particularly risk taken in peer groups) to be more alluring (Steinberg, 2007). It is the development of the planning, thinking ahead, and self-regulation networks that allows the balance between socio-emotional reward and logical reasoning to happen.

Thus, is not *either* emotional reward *or* calculated caution but an appropriate balance between them. This is reminiscent of the admonition in Aristotle (340 BCE/1941, 1106a31–1106b8) that virtue is found in the median, but that the median can be different for different people and in different situations.²⁰

2.6.3 Blending Reason and Emotion

Earlier work in moral neuroscience (Greene et al., 2001) has characterized emotion and reason as playing competitive roles in determining moral judgment and action. But we do not need to construct the relation between emotion and reason as competitive. The work reviewed earlier by Shenhav and Greene (2010) and Steinberg (2007) can also be construed as detailing how emotion/reward/escape and reason/planning/cognitive control are appropriately blended together to influence judgment and action. This is also the approach that Beer et al. (2006) take in their work. Complex moral motivations and values are represented as configurations of declarative content and emotion. Moll et al. (2003, 2009) emphasize the cooperative interaction of cognition and emotion in their model of moral judgment and track the various neural components underlying each. Moll et al. (2006) demonstrated that different cognitive-emotional complexes (considered at the psychological level) are selectively activated (measured as fMRI activation in areas associated with emotion or reason) in decisions to donate to charities with which one agrees or disagrees (e.g. regarding abortion). Thus, cultural knowledge, self-knowledge, and emotion (e.g. anticipatory guilt vs. anticipatory disgust) combine into “cognitive emotional complexes” that are balanced or play off against one another.²¹ Of course, reason and emotion can also be blended through the feedback loops associated with what has been called “embodied emotion,” the “perceptual, somatovisceral, and motoric” aspects of the ‘feelyness’ of an emotion that are perceived and then integrated into moral judgment and action” (Niedenthal, 2007, p. 1002). Ellemers and colleagues also emphasize the interaction and balance between reason and emotion (Ellemers et al., 2019; Ellemers & van Nunspeet, 2020). In Cohen’s (2005, p. 4) words, “most evaluations – from the most primitive to the most sophisticated – engage multiple different brain systems” (see also Cushman, Young, & Greene, 2010; Pessoa, 2008).

2.6.4 Planning for Cognitive-Emotional Effectiveness

How do we attempt to make these balances appropriate to the situation and to our abilities? We have some knowledge of the neural systems that help internally in making this balance, such as the orbitofrontal cortex (Beer et al., 2006). And it seems that generating options and choosing among them are separate actions with somewhat separate neural underpinning (Morris et al., 2021), with option generation being constrained by “cached value estimates” based on past experience.

²⁰ Aristotle explicitly includes appropriate emotion in his description of the characteristics of a virtuous act.

²¹ Note that this balance of complex influences is some justification for the complex model of consistency that we present in Chapter 5, and for the idea of “reflective socio-emotional reflective equilibrium” presented in Chapters 8 and 9.

Cohen (2005) argues that we use training programs and other social structures to give us the appropriate experiences (and thus cached value estimates) that help to shape our decision-making. That is, we structure our environments to support the balancing of influences as we learn to make decisions. Thus, the causal sequence that allows for balance runs *outside the brain and the person*, into the situation and culture:

For example, the specialized training given to doctors and soldiers involves the cultivation of mechanisms for averting or overcoming strong emotional responses that may interfere with their professional functions.²² These mechanisms may not rely directly on the prefrontal cortex; instead, they may involve the training of other lower-level mechanisms specific to the particular circumstances involved²³ (Cohen, 2005, p. 19)

Some of the processes to which Cohen refers here are social structures like training programs that “overtrain” particular responses such as procedures in surgery, rescue, or battle. But someone must plan the procedures and assess them. Cohen continues: “Importantly, however, the social structures that devised and support the training procedures almost certainly did rely on the prefrontal cortex” (Cohen, 2005, p. 19). These planned social structures then influence the individual as he or she performs the planned action. This planning ahead to shape our future responses has been called “situational self-control” and has been shown to be quite effective (Duckworth, Gendler, & Gross, 2016).

Thus, reasoning can be seen as something that helps to construct the future environment in order to train and to circumscribe future cognitive-emotional processing.²⁴ Oddly enough, this cultural emphasis of reason over emotion may itself be an achievement of reason in the service of training ourselves to better balance cognition and emotion in moral action.

2.7 Moral Emotion

2.7.1 Emotion as Bias

One of the main reasons that emotion has been seen as the enemy of thoughtful moral judgment is that an emotion-free standard has been assumed by early theories of moral judgment (Lapsley & Narvaez, 2005), though, of course, similar assumptions

²²Cold scientific rationality can, on the other hand, sabotage the physician–patient relationship. The technically competent doctor may appear to lack “respect and empathy” for the patient if they do not address the existential threat that the patient is experiencing (Agledahl et al., 2011). There is a new movement in physician training that focuses on integrating concern for existential and spiritual aspects in physician–patient interactions (Paal, Helo, & Frick, 2015). Again, emotion and reason regulate each other.

²³We disagree with Cohen’s implication that this sort of planning is needed to circumscribe *only* those problems associated with emotion. As we explain in Chapter 7, we should also beware of the significant failings associated with reason.

²⁴One pathology of this balance can be seen in post-traumatic stress disorder (PTSD) symptoms of dissociation of emotion and reason with regard to particular domains associated with the traumatic experience (Butollo et al., 2015).

reach back to the pre-Socratic philosophers. And emotion can indeed bias evaluation. Particularly when we are highly partisan, the way we evaluate information is driven by our emotional opposition (or commitment) to our partisan positions. Westen et al. (2006) asked political partisans to judge information that was threatening to their preferred candidate in a presidential election. They documented the neural underpinnings of a two-stage process: first an emotional response to the negative information, activating the mPFC and vmPFC, followed by a reward system activation (caudate nucleus) as soon as they had assimilated and explained away the threat.

Thus, emotion seems to be at least part of the “push” associated with implicit evaluation. Luo et al. (2006) have tracked the neural correlates of emotional engagement in what has been called the implicit association test (IAT). This involves pairing the target judgment (in this case pictures of legal or illegal behavior) with emotionally congruent pictures (positive valence animals, e.g. puppies) or incongruent (negative valence animals, e.g. snakes). Participants were then asked to make a judgment about the behavior. Reaction times for these judgments were slower on incongruent trials than on congruent ones. This suggests that the emotional incongruence is interfering with the response. Participants are not aware of this difference and even when instructed and given practice, cannot easily influence it (Steffens, 2004). Luo et al. (2006) found that amygdala response correlated with the intensity of the stimulus and of the IAT effect. This suggests that it is the amygdala that is involved in the emotional “push” to the implicit judgment that is tracked by the IAT, presumably as part of some larger evaluative system (Barrett, 2017).

These emotion-laden judgments can also drive our desire to punish others (or at least those who deserve it). Though we have a range of rational theories available to explain why punishment is a good thing to do (e.g. deterrence), in the end we punish people (or judge other’s punishment to be good) because it feels good to us. The perception of unfairness activates emotion centers such as the anterior insula (Sanfey et al., 2003) and administering punishment to an individual who deserves it activates reward centers such as the nucleus accumbens (Strobel et al., 2011). This does not, of course, resolve ethical puzzles about justifiable punishment, but it does suggest that we find it rewarding to punish those who deserve it.

Even given the significant role that emotional reaction might play in biasing judgment, it is possible to reappraise the causes of our emotional reactions (when we are aware of them) or to attempt to suppress them. Reappraisal involves reinterpreting the emotion-inducing event as something different, while suppression involves simply attempting not to experience the emotion. Most research has shown that reappraisal is more effective in controlling emotional reactions than is repression (Quirk & Beer, 2006). These two approaches to emotional regulation have different neural components that interestingly track their different outcomes. Goldin et al. (2008) had participants watch negative emotion-arousing film clips and asked them to either reappraise the emotion or suppress it. Different parts of the PFC were involved in reappraisal and suppression. In reappraisal, these areas were active early in the clip, reduced their activity as the clip continued, and showed suppressed activity in the amygdala throughout the clip. But suppression showed PFC activity increasing over the course of the film, along with increased amygdala activity. This complex pattern shows the role that various interacting networks play in constructing our experience of emotion and its meaning (Lebois et al., 2020).

2.7.2 Questioning the Cognitive/Emotion Distinction

The complex relationship of emotion to perception, judgment, and action and the dense interconnectedness of “emotional” and “cognitive” structures in the brain has led some to call for abandonment of the cognition/emotion dichotomy, at least at the structural neural level. Pessoa (2008, 2009) provides three neuroscientific reasons why the dichotomy does not stand scrutiny. Brain areas that are often labeled as “affective” are often also involved in cognition, and similarly, areas labeled cognitive are involved in affective reactions. Finally, cognition and emotion are integrated in the brain, both behaviorally in that evaluations involve both affective and cognitive areas, and structurally in that connection maps of these areas show little separation and massive interconnectivity. (See also Barrett, 2009 for a parallel view.) This complex interaction between reason and emotion is the basis for the approach to moral emotion that we review in Chapters 7 and 8.

We reviewed in Section 2.6.3 the work of Moll and colleagues (Moll et al., 2002, 2003, 2006, 2009;) on cognitive-emotional complexes. Complex moral motivations and values are represented as configurations of declarative content and emotion (Moll et al., 2009). Thus, the cognitive-emotional complexes for attachment/compassion and for anger/disgust may be separable from each other, but they are not easily separable into emotional and cognitive components. This seems congruent with the difficulty we have in separating emotion from reason in our decision-making. Perhaps we can conclude that it is not the primacy or autonomy of reason that drives human morality and makes humans unique among species but the unique ability to link emotional and motivational states to abstract ideas. It is reason *and* emotion, as they interact to construct our evaluations and our motivations toward action (Barrett, 2013; Lebois et al., 2020; Lindquist et al., 2012). And here lies one central challenge of moral acting.

2.8 Moral Formation

Chapter 9 discusses in depth how we learn skills and adopt values in order to become more moral. Part of this process involves training our self-regulation skills. As we have outlined it in Chapter 9, performing and planning is about both short- and long-term self-regulation. By self-regulation we do not mean simply self-control (suppression or distraction or reappraisal of incompatible responses to a goal) but all the goal-directed activity that is relevant to the moral life. This includes metacognition, self-reflection, and moral creativity in planning at both projects and a life-goals levels. It also includes the planning and execution of behavioral initiation, maintenance, and habit (Rothman et al., 2011), and what the philosopher Frankfurt (1971) has called second-order volitions, or “wanting to have (or not have) certain desires and motives” (p. 7) (see also Lapsley, 2008). This also includes how the stories we construct about ourselves affect and guide our moral identity and action (McAdams, 2009).

Almost all of these psychological processes are highly abstract, involving complex metacognition and self-reflection, etc., which makes it difficult to track them with current fMRI or other neuroscience methods (Barrett, 2009; Pessoa, 2008). Though we can measure the extent of neural activity in a place, we can only “read its content”

to the extent that changes in the task we give a subject will produce changes in localization. And this only allows us to infer that something about the difference in conditions produced the change in localization.

Thus, a great deal of work has been done about the neural underpinnings of emotional regulation (Damasio, 1994; Heatherton, 2011; McRae, Ochsner, & Gross, 2011; Quirk & Beer, 2006) but very little about the higher-order concepts listed earlier (Heatherton, 2011). For the rare exception to this rule, see the work of Eddington et al. (2007); and Strauman et al. (2012), which not surprisingly implicates the PFC in tracking aspirational goals. Cohen (2005) observes that we use reason (and the PFC) to construct the environment in order to train the self, but beyond that useful observation he provides no suggestion about how one might track the neural underpinnings of this process. One might expect that the affective-cognitive complexes that Moll et al. (2009) identify may play a significant role in this process, and the decision paradigm they use to track the neural correlates of charitable giving over short periods of time could provide a template for work in long-term self-regulation. Alternatively, the work on the development of cognitive control networks in adolescence (Blakemore & Choudhury, 2006; Kuhn, 2006; Steinberg, 2007) might provide another paradigm for beginning research in this area. Finally, the constructivist understanding of how situated emotions (e.g. social vs. physical fear) are learned may give some insight into how different framing of emotions can help to shape them and their neural underpinnings (Lebois et al., 2020). There is much systematic work to be done in this area if the measurement and methodology hurdles can be overcome.

2.8.1 Empathy

We not only think about others we also feel with and for them, one of the important markers for moral emotions. How do the feelings of others become connected to our own? A remapping of internal sense information in the insular cortex (located between the frontal and temporal lobes) has been proposed as the basis for the conscious awareness of how we feel, that is, for the “feelyness” of our emotions (Damasio, 1994). But this area where our feelings are represented to ourselves is also activated by many tasks that involve thinking about or perceiving other people (Adolphs, 2009). This parallel activation suggests that a kind of internal mapping of others’ feelings onto our own feelings helps to account for the sharing of emotions we call empathy (Adolphs, 2009). It also connects with the more complex description of emotion in Chapter 8 as involving the interpretation of perceived reality.

We have long known that people in cooperative interactions tend to synchronize their behavior patterns with each other (Bernieri & Rosenthal, 1991). Recent work has shown that this non-verbal matching is underpinned by what, as discussed in Section 2.2.4, is called BB synchrony. This synchrony is measured with sensitive EEG monitors that simultaneously collect brain activity information from two or more people. It turns out that the matching involved is not simply a matching of brain activation associated with non-verbal behavior but a matching based on interpersonal social connection. It is seen, for instance, between friends but not strangers, between parents and their children but not others’ children. For instance, in one study romantic couples showed BB synchrony during even short periods of shared gaze and

positive affect, while strangers needed much longer periods of shared gaze to establish synchrony (Kinreich et al., 2017). This synchrony occurred in the junction between the parietal and temporal lobes, an area that incorporates information from a wide range of other brain systems with functions including perception, emotion, and theory of mind. This tracking of neuro-synchrony promises to provide an important window on the neural underpinning of empathy shared sociality and to help us understand how the ways that the social world influences the functional neuroscience.

2.8.2 The Value Space of Moral Formation

In the formation chapter we track a possible three-dimensional space within which one might plot various trajectories of moral formation. The first dimension is movement from simple agency to agency in the service of others (Walker & Frimer, 2015). The second is a (likely flexible) preference for stability vs. change that might shape the way one views the others one wants to serve (Zapko-Willmes et al., 2021). And the third is a movement from more exclusive commitments to one's in-group towards a more expansive view of who is considered to have high moral priority, termed moral inclusion (Passini & Morselli, 2016). See Chapter 9, Section 9.3.1.3 for an in-depth explanation of this.

Two neuroscientific findings seem relevant here. Soutschek et al. (2016) have established the role of the posterior temporo-parietal junction²⁵ in egocentricity bias in a pro-social sharing task. When they disrupted the activity of this area, individuals shared less with out-group members. Sharing was reduced to a smaller extent with in-group members, and not at all for the self. This kind of discounting is likely associated with a more expansive view of who counts morally. Ellemers and van Nunspeet (2020) review evidence that neural areas associated with empathic concern and reward processing are involved in in-group favoritism and openness to moral feedback from out-group members, respectively. This begins to give us clues about how the empathic concern and reward processing might be involved in navigating the value space of moral formation.

2.9 Discussion

2.9.1 Conclusion

This somewhat scattered overview of the many different neural contributions to moral perception, judgment, and action has led us over much of the neural landscape, from middle-level sensory and motor areas to high-level cognition and emotion. One straightforward conclusion is that when we consider moral action broadly, there is no single "moral center" to the brain. Morality is served by processes and functions that are distributed throughout the brain (also in other parts of our physiology, and indeed in our culture). A second conclusion is that our moral action is neither cold nor hot but a blend of both reason and emotion. A third conclusion is that causal influence in

²⁵A location thought to be involved in taking the perspective of the future self.

moral action is more complicated than simple reductionist models suggest; moral action is structured by our brain's organization *and* our brain's organization is structured by our taking moral action. It is a cycle of influence.

1. *There is no moral "center" in the brain.*

Though neuroscience can help to map the "moral brain," there is no single neural structure one could identify as the center of morality. Instead, the moral brain seems to show just as rich a pattern of variety on the neural level as moral action does on the psychological and social level, including cultural influences.²⁶ Once we recast our understanding of morality as involving a broad range of influences and processes, as we do in this book, it becomes quite clear that morality is not localized but distributed throughout the body, embodied if you will, as well as embedded in culture.²⁷ It seems to be more a property of the person, like acting and reacting in the world and culture, than it is a property of specific brain areas or of specific psychological functions.

2. *There is little neurological warrant for a strong reason/emotion dichotomy.*

The neural basis for our taking moral action is embodied in the neural processes that occur across the massive interconnectivity of the sensory, reason/cognition/control, emotional, and motor areas of the brain. It is becoming increasingly more difficult to sustain the reason vs. emotion dichotomy in descriptions of moral judgment or action. What seems to explain individual variation in morality is less to be found in competing emotion-reason processes but more in a person's unique ability to mediate both using personal, social, and cultural resources. This ability, at least in adulthood, is not so much a ready-made capacity but rather a skill that can be practiced.²⁸ Extensive and complicated interconnections among "cognitive" and "emotional" centers are evident in any of the wide range of functions we review here as moral (Pessoa, 2008). This suggests we might be better off talking of affective-cognitive complexes (Moll et al., 2009) or patterns in brain states (Barrett, 2009; Pessoa, 2008), or neuro-synchrony (Kinreich et al., 2017) and encouraging interdisciplinary research into how the achievement of what we call social-emotional reflective balance is achieved and maintained. It is this regulated balance between emotion and reason, supported by social and cultural resources, that makes human morality unique (and uniquely complicated) rather than the likely unattainable and surely unwise triumph of reason over emotion.

3. *Causal models of the neuroscience of morality need to recognize that taking moral action is a property of the person-in-context.*

²⁶ See Sinnott-Armstrong (2016) for a comprehensive argument that neither morality nor its neural infrastructure are united in a single system.

²⁷ Because emotion is distributed throughout the body, so is morality through the extended nervous system, its interactions with the endocrine system, and the effects of both on the other tissues of the body (Niedenthal, 2007).

²⁸ See Chapter 9. It is misleading to call this balancing of emotion and rationality "emotional self-regulation." We do not seem to have a comparable study of the "self-regulation of rationality" (though perhaps we should) that would help in establishing this balance. As we see in Chapter 7, rationality can also go horribly astray, and needs somehow to be regulated, perhaps by emotions like compassion.

Neuro-essentialist, reductionist accounts of morality would say that all the important action is occurring at a neuronal level (while perhaps including the rest of the body, too). The structure of our moral action is certainly shaped by neural and physiological processes, but understanding this also requires an in-depth psychological description of moral action. This description is further embedded in processes and structures beyond the person, in groups, organizations, and cultures. And all these levels interact in influencing which aspects of which neural networks will be involved in any particular aspect of a moral action. For this reason, we will need to integrate multiple narratives of moral action to help us grasp its complexity: anthropological, sociological, philosophical, and theological narratives.²⁹

These three conclusions are interdependent. The idea that moral processing is distributed across many brain areas implies the conclusion that moral action depends on a balance or blend of reason and emotion in order to translate one's moral convictions or moral identity into a skilled action in a certain situation, context, and time. The distribution of moral processing is also at least intertwined with the idea that causality can run "outside the brain," through psychological concepts and plans, to social structures, then return as situational influence on brain processes and even structure. And the conclusion that there is a cycle of influence suggests several ways the balance between reason and emotion might be achieved.

2.9.2 Application

We began this chapter with a sketch of Henry Molaison's contributions to psychology and neuroscience and will end by applying some of the lessons we have learned in this chapter to understand him. While most of the scholarly conversation about H. M.'s case was centered on his memory, very little of the copious research on his case could have been done if H. M. had lost his social sensitivity. In the New York Times obituary, H. M. was described as unfailingly polite and welcoming to the troops of researchers who constantly visited him to ask him questions and give him tasks. This positive demeanor was consistent with his earlier character. One might describe this characteristic way of behaving to all as showing evidence of the moral virtue of hospitality.³⁰ This is true despite everyone who entered his house after his

²⁹For an example of a theological narrative, the heart has been a metaphor for the center of the person at least since early Hebrew scriptures were written (Peters, 1999). The theologian Kierkegaard does probably the earliest modern psychological analysis of morality using this heart metaphor. He describes in some complexity how an individual appropriates knowledge and values, taking them "to heart." See Chapter 9 for the complexities of this process. This analysis has influenced theologians such as Tillich and, through Tillich, psychologists such as Rollo May and Irvin Yalom (see Furchert, 2011 for an intellectual history of these connections). The neuroscience of cognitive emotional complexes gives us a different view of how this is done. But the subtlety of Kierkegaard's analysis offers different food for thought regarding the process – food that could result in a harvest of even more understanding (see e.g. Koole, Greenberg, & Pyszczynski, 2006).

³⁰For those who doubt this is a virtue, it is seen as central in the Abrahamic traditions (e.g. in Christian Benedictine monastic tradition (Merton, 2009) welcoming the stranger is akin to welcoming Christ).

surgery being a stranger of sorts to him and so needing to be welcomed anew. His consistent welcome did much for the cause of neuroscience but still cannot be laid entirely at the door of an internal virtue of hospitality. His living arrangement was such that he was constantly attended to by carers and his surrounding kept as similar as possible (given his severe memory problems). It is likely that the research visits were framed by his attendants as something to welcome, and there was a great deal of space in his schedule and support in his environment for the visit. One might say that Henry Molaison was a hospitable person in an environment that leant itself to hospitality.

2.9.3 Open Questions

1. *How are emotion and reason integrated?*

Rather than try to answer the question of emotional regulation we should be pursuing the problem of the integration of reason and emotion. We should explore models of both emotional and rational regulation. The two processes are highly interdependent, and clearly influence each other. But how does emotion help to regulate the excesses of reason? How this happens will clearly change over the life span. We have the most information on the neural underpinning for childhood and adolescence, but new methods (see Section 2.8.3.3) should help us ask some of these questions for adulthood.

2. *How does the causal cycle in moral action run “outside the brain”?*

This book’s emphasis on moral action makes us less likely to be satisfied with descriptions of morality that stay only in the head. How do social and cultural systems integrate into the cycle of influence shaping moral action? And how does personal initiative (e.g. metacognition, self-reflection, and moral creativity in planning) mediate this cycle of influence? What neural systems interact in this cycle of influence and what roles do they play?

3. *Connecting social processes and neural processes.*

We are no longer limited to suggestions that activation in particular brain regions are associated with thinking about or feeling something. Recent work in neuroscience that connects social processes with their neural embodiment are a major step forward (de Oliveira-Souza et al., 2016; Kinreich et al., 2017; Lebois et al., 2020; Tamir & Thornton, 2018). These approaches allow us to track both social processes (e.g. moral decision-making, person perception, empathy, social connection) and the shifting neural processes that underlie them over time and compare these across multiple individuals (or even with individuals in interaction with each other). Careful work of this sort can begin to give us an understanding of how the psychological processes of morality are interdependent with neural processes.

2.10 Further Readings

These suggested reading are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Barrett (2021). “This is how your brain makes your mind.” This is a readable review of the constructionist take on how our experiences, emotions, decisions, and actions are constructed by the brain and situated in situation and culture.
- Brethel-Haurwitz et al. (2018). “Extraordinary altruists exhibit enhanced self-other overlap in neural responses to distress.” A central paper in the ongoing research program by Abigail Marsh and colleagues on the neuroscience of altruism. Note the interactions between psychological terms (self-other overlap) and neuroscience findings.
- Heatherton (2011). “Neuroscience of Self and Self-Regulation.” An expansive and informative review of self and self-regulation and its neural underpinnings.
- Kihlstrom (2007). “Does neuroscience constrain social-psychological theory?” A useful discussion of the relationship between psychological claims and theory and neuroscience findings. Kihlstrom makes the case that while neuroscience is helpful in informing psychological theory and sparking creativity, it does not constrain it in any reductive sense. But see the paper by Tamir and Thornton (2018) later for an example of the two working very closely in tandem.
- Kitayama et al. (2019). “Cultural Neuroscience.” A review of the rapidly expanding field by one of its founders.
- Tamir and Thornton (2018). “Modeling the Predictive Social Mind.” Example of the sort of complex work that can be done with current methods. It connects a multilayered framework of social cognition involved in thinking about others to the neural substrate, using fMRI to show the dimensional structure of the social cognition.
- Tang et al. (2015). “The neuroscience of mindfulness meditation.” An excellent review of changes in brain structure associated with mindfulness meditation and of the psychological processes (e.g. attention, self-regulation) involved.

References

- Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual Review of Psychology*, 60, 693–716.
- Agledahl, K. M., Gulbrandsen, P., Forde, R. et al. (2011). Courteous but not curious: how doctors’ politeness masks their existential neglect. A qualitative study of video-recorded patient consultations. *Journal of Medical Ethics*, 37(11), 650–654. doi:10.1136/jmc.2010.041988
- Alvarez, G. A., Quintana, D. S., & Whitehouse, A. (2016). Beyond the hype and hope: Critical considerations for intranasal research in autism spectrum disorder. *Autism Research*, 10(1), 25–41. doi:10.1002/aur.1692
- Ambady, N., & Bharucha, J. (2009). Culture and the brain. *Current Directions in Psychological Science*, 18(6), 342–345.
- Aristotle. (340 BC/1941). *Nicomachean ethics* (W. D. Ross, Trans.). In R. McKeon (Ed.), *The Basic Works of Aristotle* (pp. 927–1112). New York: Random House.
- Barrett, L. F. (2009). The future of psychology: connecting mind to brain. *Perspectives on Psychological Science*, 4(4), 326–339. doi:10.1111/j.1745-6924.2009.01134.x
- Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review*, 5(4), 379–389. doi:10.1177/1754073913489753
- Barrett, L. F. (2017). *How Emotions are Made: The Secret Life of the Brain*. Boston: Mariner Books.

- Barrett, L. F. (2021). This is how your brain makes your mind. *MIT Technology Review* (August). Retrieved from <https://www.technologyreview.com/2021/08/25/1031432/what-is-mind-brain-body-connection/>
- Bartz, J. A. (2016). Oxytocin and the pharmacological dissection of affiliation. *Current Directions in Psychological Science*, 25(2), 104–110. doi:10.1177/0963721415626678
- Bartz, J. A., Lydon, J. E., Kolevzon, A. et al. (2015). Differential effects of oxytocin on agency and communion for anxiously and avoidantly attached individuals. *Psychological Science*, 26(8), 1177–1186. doi:10.1177/0956797615580279.
- Basso, J. C., Satyal, M. K., & Rugh, R. (2020). Dance on the brain: Enhancing intra- and inter-brain synchrony. *Frontiers in Human Neuroscience*, 14, 584312. doi:10.3389/fnhum.2020.584312
- Beer, J. S. (2007). The default self: feeling good or being right? *Trends in Cognitive Sciences*, 11(5), 187–189. doi:10.1016/j.tics.2007.01.006
- Beer, J. S., Knight, R. T., & D’Esposito, M. D. (2006). Controlling the Integration of emotion and cognition: The role of frontal cortex in distinguishing helpful from hurtful emotional information. *Psychological Science*, 17(5), 448–453.
- Beer, J. S., Lombardo, M. V., & Bhanji, J. P. (2010). Roles of medial prefrontal cortex and orbitofrontal cortex in self-evaluation. *Journal of Cognitive Neuroscience*, 22(9), 2108–2119. doi:10.1162/jocn.2009.21359
- Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rime (Eds.), *Studies in Emotion and Social Interaction: Fundamental of Nonverbal Behavior* (pp. 401–432). New York: Cambridge University Press.
- Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, 47(3–4), 296–312. doi:10.1111/j.1469-7610.2006.01611.x
- Boyke, J., Driemeyer, J., Gaser, C. et al. (2008). Training-Induced brain structure changes in the elderly. *Journal of Neuroscience*, 28(28), 7031–7035. doi:10.1523/jneurosci.0742-08.2008
- Brethel-Haurwitz, K. M., Cardinale, E. M., Vekaria, K. M. et al. (2018). Extraordinary altruists exhibit enhanced self-other overlap in neural responses to distress. *Psychological science*, 29(10), 1631–1641. doi:10.1177/0956797618779590
- Brigandt, I., & Love, A. (2017). Reductionism in Biology. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University Press.
- Butollo, W., Karl, R., Konig, J. et al. (2015). A randomized controlled clinical trial of dialogical exposure therapy versus cognitive processing therapy for adult outpatients suffering from PTSD after type i trauma in adulthood. *Psychotherapy and Psychosomatics*, 85(1), 16–26. doi:10.1159/000440726
- Campbell, M. E., & Cunnington, R. (2017). More than an imitation game: Top-down modulation of the human mirror system. *Neuroscience and Biobehavioral Reviews*, 75, 195–202. doi:10.1016/j.neubiorev.2017.01.035
- Carey, B. (2008, December 5). H. M., an unforgettable amnesiac, dies at 82. *New York Times*, p. A1.
- Casebeer, W. D. (2003). Moral cognition and its neural constituents. *Nature Reviews Neuroscience*, 4(10), 840–847.
- Cohen, J. D. (2005). The vulcanization of the human brain: A neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives*, 19(4), 3–24.
- Cushman, F., Young, L., & Greene, J. D. (2010). Our Multi-system moral psychology: towards a consensus view. In J. M. Doris (Ed.), *The Moral Psychology Handbook* (pp. 47–71). Oxford: Oxford University Press.
- Damasio, A. (1994). *Descartes’ Error: Emotion, Reason, and the Human Brain*. New York: Grosset/Putnam.

- Davis, M. H., Mitchell, K. V., Hall, J. A. et al. (1999). Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of Personality*, 67(3), 469–503.
- de Oliveira-Souza, R., Zahn, R., & Moll, J. (2016). The neuropsychiatry of moral cognition and social conduct. In S. M. Liao (Ed.), *Moral Brains: The Neuroscience of Morality* (pp. 203–236). New York: Oxford University Press.
- de Waal, F. B. M. (2006). Morally Evolved: Primate social instincts, human morality, and the rise and fall of “veneer theory.” In F. B. M. de Waal, S. Macedo, & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 1–58). Princeton, NJ: Princeton University Press.
- DeYoung, C. G., & Gray, J. R. (2009). Personality neuroscience: Explaining individual differences in affect, behavior, and cognition. In P. J. Corr & G. Matthews (Eds.), *The Cambridge Handbook of Personality Psychology* (pp. 323–346). Cambridge: Cambridge University Press.
- DeYoung, C. G., Hirsh, J. B., Shane, M. S. et al. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science*, 21(6), 820–828.
- di Pellegrino, G., Fadiga, L., Fogassi, L. et al. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91(1), 176–180.
- Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational strategies for self-control. *Perspectives on Psychological Science*, 11(1), 35–55. doi:10.1177/1745691615623247
- Eddington, K. M., Dolcos, F., Cabeza, R. et al. (2007). Neural correlates of promotion and prevention goal activation: An fMRI study using an idiographic approach. *Journal of Cognitive Neuroscience*, 19(7), 1152–1162.
- Ellemers, N., van der Toorn, J., Paunov, Y. et al. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*, 23(4) 332–366. doi:doi.org/10.1177/1088868318811
- Ellemers, N., & van Nunspeet, F. (2020). Neuroscience and the social origins of moral behavior: How neural underpinnings of social categorization and conformity affect everyday moral and immoral behavior. *Current Directions in Psychological Science*, 29(5), 513–520. doi:10.1177/0963721420951584
- Elliot, M., Knodt, A. R., Ireland, D. et al. (2020). What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychological Science*, 31(7), 792–806. doi:10.1177/0956797620916786
- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1), 5–20.
- Furchert, A. (2011). Irvin Yalom: The “Throw-ins” of psychotherapy. In J. Stewart (Ed.), *Kierkegaard's Influence on the Social Sciences* (pp. 273–295). Farnham: Ashgate.
- Goldin, P. R., McRae, K., Ramel, W. et al. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, 63(6), 577–586.
- Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, 6(12), 517–523. Retrieved from http://www.sciencedirect.com/science?_ob=MIimg&_imagekey=B6VH9-47C8S3T-H-5&_cdi=6061&_user=5671705&_orig=search&_coverDate=12%2F01%2F2002&_sk=999939987&view=c&wchp=dGLbVlb-zSkzk&md5=7686512b094c4e2bdce7bfef499c95ef&ic=/sdarticle.pdf
- Greene, J. D., Sommerville, R. B., Nystrom, L. E. et al. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105.
- Grigaityte, K., & Iacoboni, M. (2016). Merged minds: Integration of bottom-up and top-down processes for social interactions. In S. O. Sukhinder & E. S. Cross (Eds.), *Shared Representations: Sensorimotor Foundations of Social Life* (pp. 22–37). Cambridge: Cambridge University Press.

- Haidt, J. (2010). Moral psychology must not be based on faith and hope: commentary on Narvaez (2010). *Perspectives on Psychological Science*, 5(2), 182.
- Heatherington, T. F. (2011). Neuroscience of self and self-regulation. *Annual Review of Psychology*, 62(1), 363–390. doi:10.1146/annurev.psych.121208.131616
- Herdener, M., Esposito, F., di Salle, F. et al. (2010). Musical Training induces functional plasticity in human hippocampus. *Journal of Neuroscience*, 30(4), 1377–1384. doi:10.1523/jneurosci.4513-09.2010
- Heyes, C., & Catmur, C. (2022). What happened to mirror neurons? *Perspectives on Psychological Science*, 17(1), 153–168. doi:10.1177/1745691621990638
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52(12), 1280–1300. doi:10.1037/0003-066x.52.12.1280
- Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, 60(1), 653–670. doi:10.1146/annurev.psych.60.110707.163604
- Kihlstrom, J. F. (2007, 05/05/2007). Does neuroscience constrain social-psychological theory? Retrieved from <http://socrates.berkeley.edu/~kihlstrm/SPSPDialogue06.htm>
- Kinreich, S., Djalovski, A., Kraus, L. et al. (2017). Brain-to-brain synchrony during naturalistic social interactions. *Scientific Reports*, 7(1). doi:10.1038/s41598-017-17339-5
- Kitayama, S., Varnum, M. E. W., & Salvador, C. E. (2019). Cultural neuroscience. In D. Cohen & S. Kitayama (Eds.), *Handbook of Cultural Psychology* (2nd ed.). New York: Guilford Press.
- Kleber, B., Veit, R., Birbaumer, N. et al. (2009). The brain of opera singers: Experience-dependent changes in functional activation. *Cerebral Cortex*, 20(5), 1144–1152.
- Koole, S. L., Greenberg, J., & Pyszczynski, T. (2006). Introducing science to the psychology of the soul: Experimental existential psychology. *Current Directions in Psychological Science*, 15(5), 212–216. doi:10.1111/j.1467-8721.2006.00438.x
- Krys, K., Vignoles, V. L., De Almeida, I., et al. (2022). Outside the “cultural binary”: Understanding why Latin American collectivist societies foster independent selves. *Perspectives on Psychological Science*, 17(4), 1166–1187. doi:10.1177/17456916211029632
- Kuhn, D. (2006). Do cognitive changes accompany developments in the adolescent brain? *Perspectives on Psychological Science*, 1(1), 59–67.
- Lapsley, D. K. (2008). Moral self-identity as the aim of education. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of Moral and Character Education* (pp. 30–52). New York: Routledge.
- Lapsley, D. K., & Narvaez, D. (2005). Moral psychology at the crossroads. In D. K. Lapsley & F. C. Power (Eds.), *Character Psychology and Character Education* (pp. 18–35). South Bend, IN: University of Notre Dame Press.
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58, 317–344. doi:10.1146/annurev.psych.58.110405.085658
- Lebois, L. A. M., Wilson-Mendenhall, C. D., Simmons, W. K. et al. (2020). Learning situated emotions. *Neuropsychologia*, 145, 106637. doi:10.1016/j.neuropsychologia.2018.01.008
- Lindquist, K. A., Wager, T. D., Kober, H. et al. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(3), 121–143. doi:10.1017/S0140525X11000446
- Luo, Q., Nakic, M., Wheatley, T. et al. (2006). The neural basis of implicit moral attitude – an IAT study using event-related fMRI. *Neuroimage*, 30(4), 1449–1457. doi:10.1016/j.neuroimage.2005.11.005
- Lutz, A., Greischar, L. L., Perlman, D. M. et al. (2009). BOLD signal in insula is differentially related to cardiac function during compassion meditation in experts vs. novices. *Neuroimage*, 47(3), 1038–1046. doi:10.1016/j.neuroimage.2009.04.081
- Markett, S., Montag, C., & Reuter, M. (2018). Network neuroscience and personality. *Personality Neuroscience*, 5(3), 631–645. doi:10.1017/pen.2018.12
- McAdams, D. P. (2009). The moral personality. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 11–29). New York: Cambridge University Press.

- McAdams, D. P., & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology*, 61, 517–542. doi:10.1146/annurev.psych.093008.100507
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509–516.
- McRae, K., Ochsner, K. N., & Gross, J. J. (2011). The reason in passion: A social cognitive neuroscience approach to emotion regulation. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of Self-Regulation: Research Theory, and Applications* (pp. 186–203). New York: Guilford Press.
- Mendez, M. F. (2009). The neurobiology of moral behavior: review and neuropsychiatric implications. *CNS Spectrums*, 14(11), 608–620.
- Merton, T. (2009). *The Rule of Saint Benedict: Initiation into the Monastic Tradition 4*. Collegeville, MN: Liturgical Press.
- Mierop, A., Mikolajczak, M., Stahl, C. et al. (2020). How can intranasal oxytocin research be trusted? A systematic review of the interactive effects of intranasal oxytocin on psychosocial outcomes. *Perspectives on Psychological Science*, 15(5), 1228–1242. doi:10.1177/1745691620921525
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage*, 28(4), 757–762. doi:10.1016/j.neuroimage.2005.03.011
- Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: A working model. *Neuroreport*, 14(3), 299–305. doi:10.1097/OL.wnr.0000057866.05120.28
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J. et al. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *Journal of Neuroscience*, 22(7), 2730–2736.
- Moll, J., de Oliveira-Souza, R., & Zahn, R. (2009). Neuroscience and morality: Moral judgments, sentiments, and values. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 106–135). New York: Cambridge University Press.
- Moll, J., Krueger, F., Zahn, R. et al. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences*, 103(42), 15623–15628.
- Morris, A., Phillips, J., Huang, K. et al. (2021). Generating options and choosing between them depend on distinct forms of value representation. *Psychological Science*, 32(11), 1731–1746. doi:10.1177/09567976211005702
- Narvaez, D. (2008). Triune ethics: The neurobiological roots of our multiple moralities. *New Ideas in Psychology*, 26(1), 95–119. doi:10.1016/j.newideapsych.2007.07.008
- Narvaez, D. (2010). Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Narvaez, D., & Vaydich, J. L. (2008). Moral development and behaviour under the spotlight of the neurobiological sciences. *Journal of Moral Education*, 37(3), 289–312. doi:10.1080/03057240802227478
- Niedenthal, P. M. (2007). Embodying emotion. *Science*, 316(5827), 1002–1005.
- Paal, P., Helo, Y., & Frick, E. (2015). Spiritual care training provided to healthcare professionals: A systematic review. *Journal of Pastoral Care & Counseling*, 69(1), 19–30. doi:10.1177/1542305015572955.
- Pan, Y., Novembre, G., & Olsson, A. (2021). The interpersonal neuroscience of social learning. *Perspectives on Psychological Science*, 17(3), 680–695. doi:10.1177/17456916211008429.
- Passini, S., & Morselli, D. (2016). Construction and validation of the moral inclusion/exclusion of other groups (MIEG) scale. *Social Indicators Research*, 134(3), 1195–1213. doi:10.1007/s11205-016-1458-3

- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148–158.
- Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, 13(4), 160–166. doi:10.1016/j.tics.2009.01.006
- Pessoa, L. (2021). Complex systems: The science of interacting parts. Retrieved from <https://osf.io/635tq/download>
- Pessoa, L., Medina, L., & Desfilis, E. (2022). Refocusing neuroscience: moving away from mental categories and towards complex behaviours. *Philosophical Transactions of the Royal Society B*, 377(1844), 20200534. doi:10.1098/rstb.2020.0534
- Peters, J. D. (1999). Bowels of mercy. *Brigham Young University Studies*, 38(4), 27–41.
- Poldrack, R. A., & Mumford, J. A. (2009). Independence in ROI analysis: where is the voodoo? *Social Cognitive and Affective Neuroscience*, 4(2), 208–213. doi:10.1093/scan/nsp011
- Postle, B. R. (2016). How does the brain keep information “in mind”? *Current Directions in Psychological Science*, 25(3), 151–156. doi:10.1177/0963721416643063
- Quirk, G. J., & Beer, J. S. (2006). Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Current Opinion in Neurobiology*, 16(6), 723–727. doi:10.1016/j.conb.2006.07.004
- Racine, E., Bar-Ilan, O., & Illes, J. (2005). fMRI in the public eye. *Nature Reviews Neuroscience*, 6, 159–164. Retrieved from <http://pubmedcentralcanada.ca/picrender.cgi?artid=1489334&blobtype=pdf>
- Rothman, A. J., Baldwin, A. S., Hertel, A. W. et al. (2011). Self-regulation and behavior change: Disentangling behavioral initiation and behavioral maintenance. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of Self-Regulation: Research, Theory and Application* (pp. 106–122). New York: Guilford Press.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A. et al. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755–1758.
- Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, 67(4), 667–677. doi:10.1016/j.neuron.2010.07.020
- Sinnott-Armstrong, W. (2016). The disunity of morality. In S. M. Liao (Ed.), *Moral Brains: The Neuroscience of Morality* (pp. 331–353). New York: Oxford University Press.
- Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, 9(8), 1007–1008. doi:10.1038/nn1728.
- Soutschek, A., Ruff, C. C., Strombach, T. et al. (2016). Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Science Advances*, 2(10), e1600992. doi:10.1126/sciadv.1600992.
- Steffens, M. C. (2004). Is the implicit association test immune to faking? *Experimental Psychology* (formerly *Zeitschrift für Experimentelle Psychologie*), 51(3), 165–179. doi:10.1027/1618-3169.51.3.165
- Steinberg, L. (2007). Risk taking in adolescence. *Current Directions in Psychological Science*, 16(2), 55.
- Strauman, T. J., Detloff, A. M., Sestokas, R. et al. (2012). What shall I be, what must I be: neural correlates of personal goal activation. *Frontiers in Integrative Neuroscience*, 6, article 123. doi:10.3389/fnint.2012.00123.
- Strobel, A., Zimmermann, J., Schmitz, A. et al. (2011). Beyond revenge: Neural and genetic bases of altruistic punishment. *Neuroimage*, 54(1), 671–680. doi:10.1016/j.neuroimage.2010.07.051
- Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social mind. *Trends in Cognitive Sciences*, 22(3), 201–212. doi:10.1016/j.tics.2017.12.005

- Tang, Y. Y., Holzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews Neuroscience*, 16(4), 213–225. doi:10.1038/nrn3916.
- Tankersley, D., Stowe, C. J., & Huettel, S. A. (2007). Altruism is associated with an increased neural response to agency. *Nature Neuroscience*, 10(2), 150–151.
- Vul, E., Harris, C., Winkielman, P. et al. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science*, 4, 274–290.
- Walker, L. J., & Frimer, J. A. (2015). Developmental trajectories of agency and communion in moral motivation. *Merrill-Palmer Quarterly*, 61(3), 412–439.
- Westen, D., Blagov, P., S., Harenski, K. et al. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 U.S. presidential election. *Journal of Cognitive Neuroscience*, 18(11), 1947–1958.
- Wheatley, T., Kang, O., Parkinson, C. et al. (2011). From mind perception to mental connection: Synchrony as a mechanism for social understanding. *Social and Personality Psychology Compass*, 6(8), 589–606. doi:10.1111/j.1751-9004.2012.00450.x
- Wood, W., Mazar, A., & Neal, D. T. (2021). Habits and goals in human behavior: Separate but interacting systems. *Perspectives on Psychological Science*, 17(2), 590–605. doi:10.1177/1745691621994226.
- Wright, A. G. C., Creswell, K. G., Flory, J. D. et al. (2019). Neurobiological functioning and the personality-trait hierarchy: central serotonergic responsivity and the stability metatrait. *Psychological Science*, 30(10), 1413–1423. doi:10.1177/0956797619864530
- Zapko-Willmes, A., Schwartz, S. H., Richter, J. et al. (2021). Basic value orientations and moral foundations: convergent or discriminate constructs? *Journal of Research in Personality*, 92(June), 104099. doi:https://doi.org/10.1016/j.jrp.2021.104099
- Zhu, Y., Zhang, L., Fan, J. et al. (2007). Neural basis of cultural influence on self-representation. *Neuroimage*, 34(3), 1310–1316. doi:10.1016/j.neuroimage.2006.08.047
- Zuckerman, M. (1995). Good and bad humors: Biochemical bases of personality and its disorders. *Psychological science*, 6(6), 325–332.

Moral Ecology

Take my hand, my friend, we will go on together.

(Gilgamesh to his friend Enkidu: 2100 BCE/1989, p. 37)

It's everyone's fault that she don't have no money. It was stolen because people just don't care about the community ... Everybody should care that she got her money stolen

(Power, 2004, p. 51)

The two quotations that begin this chapter anchor this review in the antiquity of friendship and in the importance of community concern. The theme of friendship as a support in moral action is as old as Gilgamesh's tale from over 4,000 years ago. In the quotation, he encourages his friend Enkidu who has lost his nerve just before their epic battle with the giant Humbaba. Some 1,700 years later, Aristotle recognized the importance of friendship in supporting moral action by including a significant discussion of it in his work on ethics (Aristotle, 340 BCE/1941; Schroeder, 1992). In the second quotation, a member of a Kohlberg-style "just community" school makes an argument for collective responsibility, in which all members of a community are responsible for the moral action of the community.¹

For decades the psychology of morality has been focused on situational influences that often overwhelm any personal virtues or inclinations to moral action. In this chapter, we will broaden the focus to the moral ecology in which actors are imbedded and with which they interact. Our overview will begin with an argument for the term *moral ecology*, and then survey levels of moral ecology beginning with issues of intra- and inter-cultural influences, moving to organizational and group influences, and then to friendship. We then consider how individuals actively shape their moral ecology. We will conclude, among other things, that we can think of moral ecology as a design problem (Epley & Tannenbaum, 2017) and that the same principles that can produce a moral community can also be used to produce cultures of corruption and the social machinery of evil.

¹See Chapter 7 for a description of Kohlberg's approach to moral psychology.

3.1 How Moral Ecology Changes the Conversation

3.1.1 Why Moral Ecology?

A variety of terms have been used to identify social influence on moral action, and most seem too simplistic or limiting. *Situational influence* (Rauthmann & Sherman, 2020; Zimbardo, 2004) makes the influence sound narrowly unidirectional, in addition to limiting the focus to influence, ignoring underlying values, goods, procedures, roles, and other aspects of what Helton-Fauth et al. (2003) have called the *socio-technical system*. *Organizational culture* (Hofstede, 2001; Leidner, 2006) allows for the complexity we mention, but its disadvantage is similar; culture is at times treated as monolithic, its influence on the individual seems unidirectional, and it appears to occur only in organizations. *Organizational climate* (Denison, 1996) has the advantage of suggesting how flexible and adaptive influence can be and also encourages the analysis of macro vs. microclimates – what Brief, Buttram, and Dukerich (2001) have referred to as moral microcosms. But *climate* still shares the disadvantage of a unidirectional causality on the isolated individual.

Ernst Haeckel (1866, p. 286) first used *oecologie* to revise the unidirectional and hierarchical concept of “food chain” and to describe the interlinked, interwoven, and interdependent lives of plants and animals.² Robert Park (1952) of the Chicago school of sociology used *social ecology* to refer to the spheres and environments in which humans exist. Park viewed this complex interaction as mediated by morality, custom, and law – unlike the ecologies of all other animals.³ In philosophy, Flanagan (2008) has used *human ecology* with a similar referent to that of Park (1952).

Huff and colleagues (Huff & Barnard, 2009; Huff, Barnard, & Frey, 2008a, 2008b) use the term *moral ecology* because it encourages us to consider the complex web of relationships and influences, the long persistence of some factors and the rapid evolution of others, the variations in strength and composition over time, the microecologies that can exist within larger ones, and the multidirectional nature of causality in an ecology.

This approach casts the issue as significantly larger than the much-publicized debate over whether situational influence makes virtue impossible (Doris, 2002; Doris & Stich, 2005; Mischel, 1968, 2004; Snow, 2010).

3.1.2 Beyond Virtues vs. the Situation

First, the reader should note that the moral ecology approach we take in this text differs from the more traditional deconstructive critiques of morality in social psychology (Nisbett & Ross, 1980; Zimbardo, 2004) and associated recent philosophy (Doris, 2002; Doris & Stich, 2005). These approaches often oppose individual virtue to situational influence. They take morality to be synonymous with individual virtue,

²Haeckel’s work is in the tradition of Humboldtian science that we mention in the Introduction.

³But see Chapter 1 to see how complicated this claim is.

exercised without influence from external sources. They then show the pervasive influence of situational aspects that are often hidden from conscious awareness. Their conclusion is that true morality is either severely limited or not possible.

In contrast, a moral ecology approach sees morality as arising out of the combination of numerous interacting influences. In this complex interweaving of processes, the distinction between external and internal influences becomes difficult, as does the distinction between a “pure” moral decision and illicit biases on that decision. Moral action arises out of the interaction of the individual within the moral ecology, and multiple influences both help and hinder that moral action.

3.2 Culture and Moral Diversity

Much of the work on culture and morality has been a catalog of differences: this group in this culture thinks cutting one’s hair while in mourning is immoral, that group from another culture is puzzled by the judgment (Shweder et al., 1997, p. 131). The documentation of the diverse ways of being moral is indeed an important first step in understanding how people are moral. It has often been conceptualized in terms of evidence for better or worse morality (Kohlberg, 1963; Kohlberg & Mayer, 1972). However, more recent work has taken to simply documenting the diversity in ways of being moral (e.g. in goals, purposes, principles, skills, values, etc.). The catalog of diversity has had the salutary effect of making a plausible case that one need not posit a unitary morality, identical across cultures, and that doing so makes it more difficult to recognize the multitude of different ways of being good. In this section on culture, we will follow long tradition and give an incomplete catalog of these different ways of being good or valuing the good. At the end, we will introduce work that is asking the question of unity in the diversity: are there dimensions to the values that are similar across cultures or similarities in the processes by which culture has its influence on morality?

What has been missing until recently in the analysis of cultural differences is specification of the internal psychological processes by which these cultural differences have their effects (Heine, 2010; Markus & Kitayama, 2010). Inglehart and Welzel (2005) point to social and historical patterns that produce the differences, as does Hofstede (2001). Haidt (Haidt, 2001, 2010) has specified nonconscious social processes that underlie judgments, and we will cover these in Chapters 7 and 8. Additionally, Schwartz and colleagues (Roccas et al., 2002) have investigated the relationship between the values in Haidt’s (2001) system and personality characteristics. But this work has been somewhat scattered. Markus and colleagues (Hamedani & Markus, 2019; Markus & Kitayama, 2010) propose a model of the “cycle of mutual constitution” of cultures and selves. They specify multiple pathways through which culture influences individuals and individuals influence culture, including:

- The construction and maintenance of *culturally pervasive ideas*.
- *Institutions* that shape ideas, language, and cultural narratives, and that exert direct power (e.g. education, political and legal entities). We would also add religious (Bellah, 2011) and medical institutions (Collins & Pinch, 2006) to this list.

- Daily *situations and practices* that shape behavior and ideas at home, school, and the workplace. We would also add the built environment (Friedman & Nissenbaum, 1996; Winner, 1980).
- The construction and maintenance of self and identity in interaction with the aforementioned influences.

One can find this theme of mutual constitution throughout this chapter, as we explore how various levels of culture influence individuals and how individuals, in concert with each other, influence culture. Understanding these processes, and how they might differ in different cultures and subcultures, will be crucial to understanding how people live out their morality within their moral ecology. There is still much work to be done here.

3.2.1 Moral Diversity *Within* Cultures

One need not begin with *cross*-cultural differences in morality to catalog differences. There is sufficient moral diversity *within* cultures to provoke wonder (Heine, 2010; Kitayama et al., 2006).

3.2.1.1 Political Values A long-standing research program on moral development based in Kohlberg's stage model has identified one of these dimensions of difference: extensiveness and diversity of social experience.⁴ Within cultures, those who have access to a wider variety of social experience (as indexed by such things as age, socioeconomic status, social class, education) tend to think more that moral judgment requires shared standards of judgment, rather than simple obedience to authority.⁵ This is a difference found within each of the many cultures studied by those in the Kohlberg tradition (Gibbs et al., 2007).

The Kohlberg tradition sees these differences in moral judgment as evidence of moral development from less complex to more complex cognition. But others have cataloged differences in the reasons individuals give for moral judgment and treated them more as simple dimensions of difference. Haidt and colleagues (Graham, Haidt, & Nosek, 2009; Haidt & Joseph, 2004) have cataloged differences on five dimensions of judgment: harm/care, fairness/reciprocity, in-group/loyalty, authority/respect, and purity/sanctity. Many studies across a variety of cultures again document a difference *within* each culture: classical liberals are those who see harm/care and fairness/reciprocity as foundational values and are suspicious of the other three dimensions since they have been used unfairly to harm minorities and the disenfranchised. Classical conservatives try to balance all five dimensions and refer to all of them in their judgments. Haidt & Kesebir (2010) connect these value differences to a distinction made by Tönnies (1887/2001): *Gesellschaft*, an approach that sees a fair

⁴We list Kohlberg's approach under "political values" because commitment to the principle of justice at least overlaps with the political realm. It is no longer reasonable to think of justice as the only dimension in moral judgment, and it needs to be grouped with other similar values approaches.

⁵This corresponds to a shift from pre-conventional to conventional moral judgment in Kohlberg's approach. See Chapter 7 for a short explanation of Kohlberg's approach.

public square as essential in a diverse *society*, and *Gemeinschaft*, an approach that values trust, cooperation, and mutual aid in tight-knit *communities* (see also Markus & Kitayama, 2010 on this distinction). These differences in approach produce significant political strife and mistrust within societies that host them (Haidt & Kesebir, 2010) and correspond to two very different notions of the good. But they are, in the end, both widely accepted notions of the good within and across all cultures. And each can serve as a source of critique if the argument is reframed persuasively (Voelkel & Feinberg, 2017).

3.2.1.2 Religious Influence Work by Cohen and colleagues (Cohen & Rozin, 2001; Cohen, Rozin, & Keltner, 2004; Cohen, Siegel, & Rozin, 2003) documents the differences in moral thinking between Jews and Christian Protestants in America. The Jews they surveyed viewed practice – what one actually does – as the most central moral aspect of their religious approach, while the Protestants viewed belief and practice to be equally important. Members in the groups are defined differently (for Jews primarily by descent, for Protestants primarily by belief), and mental states such as temptation appear to matter more to Protestants than to Jews, for whom right action seems to be more important. Mormons and Muslims appear to moralize their religious norms, while Jews appear not to, suggesting there might be significant differences among religions on even what category of things one can call moral (Levine et al., 2021). Any culture that is multireligious can expect to find similar disagreements about moral issues. Fundamentalist approaches to religion can divide societies even if they are predominantly of one religion (Marty & Appleby, 1991).

Indeed, in the current political climate in the United States, differences in religiously based values are a central issue (Graham & Haidt, 2012; Haidt, 2012). There is some evidence that political and cultural differences drive the religious ones (R. P. Jones, 2016). There appear to be individuals who are primarily motivated by their “coalitional commitment” to a religious group, rather than by personal religious devotion (Ginges, Hansen, & Norenzayan, 2009). And these individuals are more likely to endorse violence against out-groups (Bloom, 2012; Ginges et al., 2009), an extreme and unfortunately common example of the pattern in certain kinds of religious/cultural commitments.

However, one can also use religious principle to *critique* a conventional morality that falls short in some respect. One of the central characteristics of religion is its critique of conventional morality (Bellah, 2011). Gregg (2005, chapter 3) reviews how the values of Islam help to counter and limit the honor/shame system in Middle Eastern countries, and even help to “set the developmental task of identity formation ... to integrate pursuits of honor and piety into a whole, or at least a balanced life” (Gregg, 2005, p. 167).⁶

Gibbs (1977) claims, for instance, that the post-conventional reasoning one can find in Kohlberg’s interviews is not always simply justice based but can be more existential- or religion-based critiques of conventional morality (Gibbs et al., 2007).

⁶Kohlberg would score basic religious commitments as a part of conventional morality (Richards & Davison, 1992) but might score more sophisticated religious reflection as post-conventional (Gibbs, 1977).

Thus, the influence of religion in culture and in morality is multifaceted, drawing from group commitments and personal piety, and providing organizational support and opportunity for action (Matsuba, Hart, & Atkins, 2007; Oliner & Oliner, 1988).

3.2.1.3 Class Differences Cultures with class differences can also expect to entertain moral diversity. In a series of experimental studies, Stellar et al. (2011) have found that lower-class Americans exhibited more compassion than upper-class individuals (e.g. in verbal response to other individuals and even in heart rate). Several studies have also shown that lower-class individuals tend to have an interdependent view of the self, while upper-class individuals value independence (Heine, 2010). Complications quickly arise in this picture since, for instance, Bowman, Kitayama and Nisbett (2009) have found that middle-class Americans gave and received more help to each other than working-class individuals, who reported more self-reliance. Like religion, we are likely to find the cultural differences embedded in class to be complex and multilayered.

3.2.1.4 Individual Variation in Goals Another kind of moral diversity is evident in the wide diversity of moral goals that individuals set for themselves. The pioneering work of Colby and Damon (1992) in moral exemplars shows great variety of goals among their sample of exemplars, even though their sample was primarily limited to individuals doing social service. These fell into two large groupings, those interested in reform of society and those interested in direct service to those in need. But even in these groupings, the exemplars they interviewed had some focus for their efforts, rather than being a “jack of all trades” in moral excellence. Huff and Barnard (2009) found a similar distinction among exemplars in the computing field, though they also found about a third of their exemplars fit neither category and instead cared deeply for some other passion, such as mentoring, education, or particular social-technical issues.

Walker, Frimer, and colleagues (Matsuba & Walker, 2004; Walker & Frimer, 2007, 2009; Walker & Hennig, 2004; Walker & Pitts, 1998) have done the most systematic work in identifying clusters of goals and personality characteristics that identify highly moral individuals. They have found clusters of “communal” and deliberative (Walker, Frimer, & Dunlop, 2010), just, brave, and caring (Walker & Hennig, 2004), and caring/dependable and principled/idealistic (Reimer, DeWitt Goudelock, & Walker, 2009). The work on exemplars is somewhat cross-cultural, including British, Scandinavian, Canadian, and American samples. Still, all these subtypes and different goals of moral excellence are from Western, educated, industrialized, rich democratic (WEIRD) societies (see Henrich, Heine, & Norenzayan, 2010) and it seems unlikely that the variety will diminish as further research encompasses other cultures.

Thus, even within cultures one can find extensive moral diversity. Most of the work that relates to within-culture moral diversity has been conducted in WEIRD societies (Henrich et al., 2010) though we are beginning to see some within Asian cultures (e.g., Kitayama et al., 2006). The variety of moral sub-ecologies within any culture is likely to be wide, including regional, class, religious, race, and organizational differences, and the interactions of these. For instance, a national concern that corporations in Japan no longer serve the “household” function they used to has resulted in a recent

upswing in interest in ethics among engineering schools and firms in that country (Downey, Lucena, & Mitcham, 2007). In contrast, because of their complicity in WWII, engineering firms in Germany have recently required a strong commitment to ethical practice, but engineering firms in France rarely talk explicitly about ethics, instead concentrating on the role of engineers as the elite vanguard of progress for the nation. In each of these cultures, local history and values have resulted in different ethical/moral approaches to a professional field. We really are just beginning to find our way in this multitude of differences. Work by Muthukrishna et al. (2020) provides a useful index of cultural distance that can be computed between any two cultures based on the World Values Survey. Dimensional models like this will facilitate comparisons across multiple cultures on psychologically relevant dimensions.

3.2.2 Patterns in Moral Diversity *Across* Cultures

There is a much longer history of research in the psychology of cross-cultural differences (see Heine, 2010 for an overview). It is from this work that we will find some pattern in the diversity of cultural difference. There is evidence for some cultural moral universals. For instance, Curry, Mullins, and Whitehouse (2019) found broad agreement across sixty societies in basic principles that allow for cooperation. And Kinnier, Kernes, and Dautheribes (2000) provide four major categories of agreement across world religions.

We have already covered the convergence across cultures of the early stages of the Kohlberg framework. Extensive work in this area (see Gibbs et al., 2007 for a review) has shown that access to broad social differences can serve as developmental pressure to move individuals from preconventional to conventional moral stages. But this unity falls apart in what Kohlberg hypothesizes to be the higher post-conventional stage of moral development – for Kohlberg, this is a movement to a concern for principles of equal justice. It turns out that different cultures have different preferences for the principles that can guide moral judgment at these higher levels. For instance, research shows that Americans prefer equity as a means of distributing scarce goods, while those from India prefer need and those from Japan prefer equality (Heine, 2010). Some theorists suggest that, rather than the “higher stages” being a Kohlbergian shift to thinking of a particular type (justice) across cultures, instead those who question the conventional judgments *within their society* find culturally embedded and supported ways of questioning, supporting, and providing flexibility to conventional wisdom (Gibbs et al., 2007). Thus, they are still “post” conventional but question and modify conventional wisdom in a way that is locally relevant and culturally supported.

On a more descriptive level, Schwartz and colleagues (Bardi & Schwartz, 2003; Roccas et al., 2002; Schwartz, 2006, 2010) have shown a two-dimensional organization that defines the value dimensions on which individuals from differing cultures disagree. Work in over seventy countries using a variety of instruments converges on this solution that represents the similarities in the way values relate to each other, and the differences in the emphasis different cultures and individuals make on those values. Thus, there is agreement about what dimensions are important but variation on which ends of the dimensions to emphasize. Figure 3.1 shows the standard circumplex of these values, with the two dimensions labeled

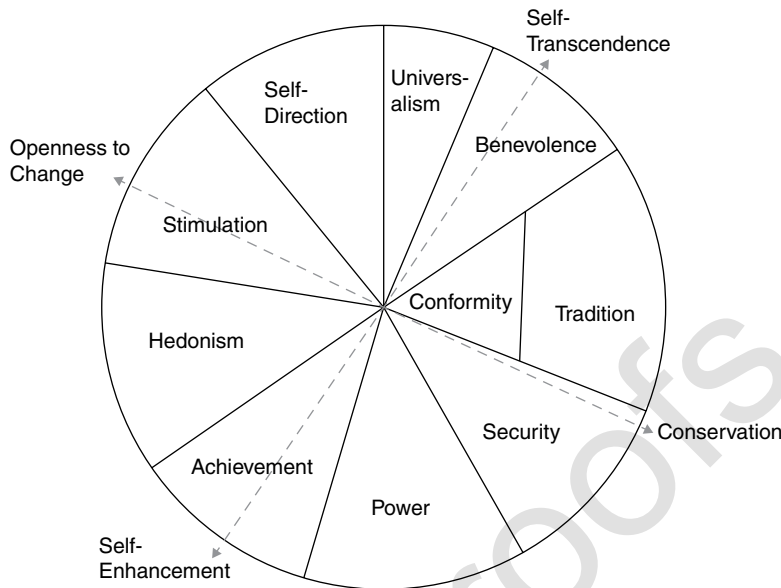


Figure 3.1 Schwartz circle of value regions.

Source: Adapted from Borg et al. (2015).

self-enhancement – self-transcendence and conservation – openness to change. Support of any one value on the circle tends to correlate with support for adjacent ones, and less support for those on the opposite pole. Work by Geert Hofstede (2001) and by Ronald Inglehart & Christian Welzel (2005) makes similar points about variation in values among cultures.

Distinctions among societies on the individualism–collectivism dimension have a long history in cultural psychology (Triandis et al., 1986) and seem also to be the most psychologically fruitful, encompassing a variety of cross-cultural differences (Heine, 2010). Individualistic cultures emphasize both the self and others as independent, agentic, and interested in self-enhancement. All these differences, and many others relevant to this dimension (see Heine, 2010 for a review) have relevance for how people from within these cultures think about individual responsibility and agency. In an interesting convergence of within- and cross-cultural patterns, Kitayama et al. (2006) have shown that when frontier conditions prevail in a subset of a collectivist country (as in the Hokkaido frontier of Japan in the late 1800s) one finds a more individualist subculture emerging. A similar historical mechanism has been theorized to be behind the honor culture in the southern United States (Nisbett & Cohen, 1996).

All the variation we have seen so far should lead us to be skeptical of work in moral psychology that makes universalistic claims but is done primarily in WEIRD cultures. When we look cross-culturally at issues of morality we find both difference and similarity. There are differences in the kinds of values or moral foundations individuals

hold (both cross-culturally and within cultures). But there are patterns to the differences. Haidt and Craig's (2004) five foundations, Schwartz's (2010) values circumplex, Hofstede's (2001) intercultural dimensions, and Inglehart and Welzel's (2005) World Values Survey all point to patterns in the ways that cultures and individuals both agree and disagree on moral issues. Recent work by Schwartz and colleagues (Zapko-Willmes et al., 2021) suggests that the value circumplex and the moral foundations approach have both a shared structure and differences. More work like this is needed to help find some convergence among the diverse systems of cross-cultural value difference.

3.3 Moral Ecology at the Organization Level

Organizations are important influences on moral action in part because of their impact on individuals working or volunteering within them and in part because of the profound influence they exert on society and culture. For instance, the Association of Certified Fraud Examiners (2018) estimated that unethical conduct in organizations cost the global economy nearly US\$4 trillion in 2017. Here we will review processes supporting both ethical and corrupt practices in organizations.

3.3.1 Organizational Values and Goals

Organizations can also provide variation in moral environment within a culture or industry. In interviews with engineers in the Chicago area, Davis (1998) has found that engineering firms differ systematically in their moral ecology, and thus in the way they treat the engineering role. *Finance-driven* organizations give priority to maximizing financial goals, and engineers in these firms do not participate in decision-making, but instead stand outside the process in a consultant-like role, providing information, producing design options, and answering questions. In this environment, engineering values are subordinated to management goals, and the engineer's ethical responsibilities are limited to exercising *due care* when providing advice and services, avoiding *conflicts of interest*, and remaining *loyal* to the legitimate interests and objectives of managers. This produces the classic *due care* blame-avoidance memo documenting concern about a product or design strategy. In the other two kinds of organization, engineers participate in decision-making, and managers seek consensus with engineers. *Quality-driven* organizations see the achievement of quality products as the defining goal. For instance, one engineer is quoted by Davis (1998, p. 133) as saying, "Cost comes in only after quality standards are met." And another offers that, "If a customer wants to take a chance, we won't go along" (Davis, 1998, p. 133). *Customer-driven* organizations see superior customer service as a defining goal. All three kinds of organization are constrained by financial issues, quality, and customer service, but they differ in terms of which goal they see as primary or fundamental. It is these differences that result in differing moral ecologies in the organizations and in differences in the moral roles and moral choices of the engineers in them, and the moral skills they need to practice.

3.3.2 Processes Supporting Organizational Corruption

One way to ask about how organizations construct moral ecologies is to look at how they construct immoral ecologies, the systematic planning and execution of illegal activities. There is a significant literature in this area. In a review of work on organizational corruption, Ashforth and Anand (2003) propose three processes by which organizational misbehavior is supported: *rationalization*, *institutionalization*, and *socialization*. In *rationalization*, actors in the organization convince themselves that the behavior is legitimate. These acts are then *institutionalized* as a matter of routine in the organization, and those new to the organization are selected and *socialized* into the routine performance of the actions. These processes provide a structure for our review of influence toward both ethical and unethical action in organizations.

3.3.2.1 Rationalization There are at least two approaches that individuals can use to separate action from its moral implications: domain limitations and cognitive distancing. *Domain limitations* are ways to make moral consideration in a particular situation or domain not relevant. Thus, the moral question simply does not arise. Once the moral question arises, then one can use a variety of *cognitive strategies to minimize* its relevance.

Moral action can seem relevant in some domains or circles of action but not in others. Colby and Damon (1992) note this in the lives of their moral exemplars: people who give their lives to causes such as racial equality or feeding the poor often treat their families with much less concern than one might expect for such compassionate individuals. In organizations, this occurs in part because of the way the organization defines the situation. Treviño, Weaver, and Reynolds (2006) speak of this as “situationally defined identities becom[ing] entrenched within organizations” (p. 962). In one study (Weber & Wasieleski, 2001) business managers enrolled in MBA courses were found to respond at lower levels of moral reasoning (in Kohlberg’s scheme) when asked about work-domain issues than asked about non-work-domain issues. The salience of the moral aspects of an action is influenced by what T. M. Jones (1991) calls moral intensity. Moral intensity is a combination of: (1) magnitude of the consequences; (2) concentration of the effect; (3) probability of the effect; (4) its temporal immediacy; (5) social consensus on the moral status of the action; and (6) proximity (see Treviño et al., 2006, pp. 953–954 for a review of research on this concept). High moral intensity can jar the senses and make it difficult to maintain a definition of the situation.⁷ Still, as the Milgram (1963, 1974) studies have shown, a definition of the situation can trap the actor into a situation even when moral intensity is quite high. It is the combination of high moral intensity (giving clearly painful shocks to a helpless, suffering, and protesting acquaintance in the next room) and the strong situational definition (having promised to be a good subject to help science) that produced the stress reactions that Milgram’s subjects experienced.

⁷One can imagine designing organizational procedures to highlight the moral intensity of proposed actions (Epley & Tannenbaum, 2017). For example, in software design, design procedures that incorporate concern for those affected by the software are available (Leveson, 1995; Nissenbaum, 2011; Shilton, 2018; Umbrello, 2019).

When domain limitations fail in making the moral aspects of organizational actions irrelevant, there are a host of cognitive strategies available to help the individual justify the course of action. In their article on corruption, Ashforth and Anand (2003, pp. 16–22) list legality, denial of responsibility, denial of injury, denial of victim status, denial of status to those who critique, appealing to higher loyalties, metaphor of ledger, and refocusing attention. Bandura (1999, 2002) has a similar list of “moral disengagement strategies” compiled from a separate theoretical approach: moral justification, advantageous comparison, euphemistic labeling, minimizing, ignoring, misconstruing, dehumanization, blaming the victim, displacement of responsibility, and diffusion of responsibility. The point of all these disengagement strategies is to distance the actor from moral responsibility for an action.⁸

3.3.2.2 Institutionalization When an action becomes “standard operating procedure” it becomes institutionalized (Ashforth & Anand, 2003). This means that it no longer requires a decision process. In institutionalizing corruption, it is usually the avoidance of a decision process that is institutionalized (Ashforth & Anand, 2003). It is not just procedures that are institutionalized, the structure of an organization can also be affected. Since the adoption of the Federal Sentencing Guidelines for Organizations in 1991 made them a factor in recommended criminal sentences, organizational ethics programs in the United States (often called offices of ethics and compliance) have become a standard part of organizational hierarchy (Stansbury & Barry, 2007). Organizations that want to provide plausible deniability for ethical misconduct have been known to establish a position to support corrupt practices ironically titled by Braithwaite (1989, p. 350) “vice president responsible for going to jail” allowing others to be willfully blind of misconduct.

In cases of corruption, the organization “comes to expect and then depend on the payoffs from corruption. In time, goals, budgets, information flows, rewards and punishments and so on may be skewed to support the practices” (Ashforth & Anand, 2003, p. 9). The common claim is that, particularly in countries known for corrupt business practice, this sort of activity is required to do business. This claim seems only to be partly true. In a study of 480 of the world’s largest corporations across many different cultures, Healy and Serafeim (2012) show that organizations that report and combat corruption have slower but more profitable sales growth and return on investment, but that this difference only occurs in high corruption countries. They speculate that establishing a reputation as corruption free gives an advantage that results in slower sales growth but increases in return business based on trustworthiness. The corruption-fighting companies establish this reputation in part by institutionalizing policies and procedures that encourage and support reporting (Healy & Serafeim, 2012).

One can institutionalize ethical action in firms by adopting business procedures that embed ethical considerations as in the proposed procedure of Helen Nissenbaum (2011) for privacy in software or that of Nancy Leveson for safety in complex systems

⁸A systematic comparison of the Ashforth and Anand (2003) and Bandura (1999, 2002) strategies with the dimensions of moral intensity would be an interesting endeavor and might bring some order to what now stand as mere lists.

(1995). Another proposal for institutionalizing ethical reflection and action is the idea that one can design this into organization structure (Epley & Tannenbaum, 2017).

A widely recognized form of institutionalization to support ethical action is the adoption of an organizational code of ethics. The *bad news* for such approaches is that across the broad range of codes of ethics in organizations, there seems to be little effect on employee or organizational misconduct (Treviño et al., 2006). The *good news* is that one approach to such codes does offer some promise: what Treviño et al. (2006) call “value internalization” approaches. These approaches are less about compelling compliance with rules and more about providing support and socialization. They operate through the organizational culture and leadership commitment, and through what Gehman, Treviño, and Garud (2013) call *values practices*: those activities that embody and shape concern for values in an organization. These practices not only shape employees but also reaffirm and maintain (or challenge and modify) organizational structure and are thus a vital part of institutionalization. An example of such practices comes from a study of the daily interactions in a national funding agency responsible for research in information and communication technology (Grimpe et al., 2020). In-depth interviews revealed that the workers collaboratively constructed the organization’s response to responsible conduct of research issues by embedding the issues in daily conversations and tasks. It became a matter of habit to take these issues into consideration in everyday tasks, and in this way, without fanfare, the workers were contributing to the moral ecology of the organization. As long as employees think basic procedural justice is being served in the organization (e.g. one does not suffer for ethical behavior), then these values practices are more broadly effective than compliance-based efforts (e.g. punishing noncompliance with rules). This pattern is called the *paradox of control* (Stansbury & Barry, 2007), where attempts to directly control employees’ ethical actions are viewed as coercive,⁹ while the socialization approach of value internalization gives flexibility and professional commitment to employees’ existing ethical goals. A widely cited form of this advantage for socialization rather than control approaches is the relative success of socialization-based (vs. control-based) diversity programs in organizations (Dobbin & Kalev, 2016; Paluck et al., 2021). This brings us then to the topic of socialization.

3.3.2.3 Socialization As implied by the paradox of control, socialization is more than simple passive absorption of the dominant culture. It involves active choice on the individual’s part. Much of the description of socialization here is taken from the organizational psychology literature, but parallels can be found in cultural psychology (Markus & Kitayama, 2010) and developmental psychology (Tomasello & Vaish, 2012).

The attraction–selection–attrition cycle (Ashforth & Anand, 2003) is the gateway to socialization in organizations. Individuals with similar values are attracted to the organization, the organization seeks out these people, and those who find a bad fit between themselves and the organization often leave. Once hired, participation in values practices helps to educate, influence, and motivate employees to adopt the

⁹See references to “Darley’s Law” that any system of coercion will be gamed (Darley, 1994) in proportion to the extent that its metrics are quantitative.

values and practices of the organization. People don't expect to suffer for ethical behavior, but the relationship between reward and ethical behavior is complex (Treviño et al., 2006). Socialization occurs primarily at the local level and, though influenced by organizational climate and ethics codes, it is identification with veterans (Ashforth & Anand, 2003) and local values practices that carry the day. For instance, the extent to which ethical language is actually used in a company is a good (though not infallible) guide to ethical behavior within it (Gehman et al., 2013). Another effective strategy is reintegrative and dis-integrative shaming (Braithwaite, 1989; Stansbury & Barry, 2007) involving public attempts to influence those who do not seem to understand the local culture (whether that culture supports or opposes corruption) with an eye to inducing them to join in the culture or to leave it. This moral seduction leads the receptive newcomer to view the world through the "moral microcosm" of the local organization unit (Moore et al., 2006).

3.4 Close Relationships in Friendship and Community

If local units within organizations are the most powerful agents of moral ecology for the individual, then outside of organizations the most powerful agents should also be local: small communities and personal relationships. In this section, we look at the most intimate and smallest scale of moral ecology.

Ironically, one of the pioneers of research on the role of the community in morality is the person most known for championing the cognitive aspects of morality, Lawrence Kohlberg. In part in reaction to his experiences on a kibbutz (Walsh, 2001), Kohlberg helped found the first *just community* school, the Cluster School in Cambridge, MA. This is the school from which the quote at the beginning of this chapter is taken. Though just community schools were designed to be democratic, with students and staff each having only one vote, their primary innovation was the attempt to develop a moral community: "a group that shares an explicit commitment to a common life characterized by norms embodying high moral ideals" (Power, 2004, p. 50). The approach was theoretically based in Durkheim's (1925/1973) work on moral communities in education.¹⁰ The just community approach recognized that norms and values were important but also emphasized individual commitment to shared ideals within the framework of a responsible community.¹¹ As Power explains (2004, p. 52): "The self does not experience a sense of obligation or responsibility to act in isolation but with others within a cultural setting." The just community schools and their progeny serve as laboratories in which to observe these processes.

But of course, these processes do not occur only in laboratories. They are crucial to the maintenance of all communities. Daniel Hart has pursued a research program on moral commitment and volunteerism in neighborhood and community that has charted the ways that individual moral commitment and development are rooted in

¹⁰This grounding in Durkheim is somewhat ironic, given Durkheim's clear relativism and Kohlberg's trenchant critique of relativism (Eberhardt, 2014).

¹¹In this respect, Kohlberg and Gilligan were much closer together in their approaches than is commonly realized.

social groups (Hart, 2005; Hart, Atkins, & Southerland, 2006; Hart & Carlo, 2005; Hart & Fegley, 1995; Hart & Matsuba, 2007, 2009; Matsuba et al., 2007). He summarizes:

Moral identity, a commitment to advance the welfare of others that is consistent with one's self-image and moral goals, emerges from the interplay of family background, personality, moral cognitions and attitudes, self-perceptions and moral emotions, and social relationships and interactions with social institutions. No single element is the keystone; there are multiple routes to moral identity formation. (Hart & Matsuba, 2009, p. 228).

However, one consequence of moral identity being embedded in community is that it then becomes subject to domain limitations. It becomes focused on our obligations to each other, often to the exclusion of others, resulting in *moral collapse* (Hart, 2005). When identity becomes fused within a limited community in this way, great evil can result (Graham & Haidt, 2012; Hart & Matsuba, 2007; Staub, 1999). Community-based morality is a fragile achievement that requires constant maintenance and is in part a result of the good fortune and thoughtful choice of one's social surroundings (Hart, 2005, p. 260).

The smallest unit of social influence in morality is the dyad. Work on the moral psychology of this level of influence has been quite rare until recently. But work on what has been called the *Michelangelo phenomenon* has important implications for thinking about how individuals achieve and maintain their moral action. The Michelangelo phenomenon occurs when one member of a dyad helps to shape a close partner's skills and traits in support of achieving the partner's ideal self (Rusbult, Finkel, & Kumashiro, 2009).¹² Research has mostly concentrated on romantic partners, though Rusbult et al. (2009) argue that the effect should equally be seen in family, friendship, and collegial relationships to the extent that these are closely interdependent. In another example of the paradox of control, this sculpting effect requires special conditions. First, the partner must be supportive of the other's *ideal* self, not just supportive of what the partner thinks is best for the other (Drigotas et al. 1999). In fact, this latter kind of "support" appears to hamper both the partner and the relationship (Rusbult et al., 2009). Second, the support for the ideal self of the other needs to be in the context of movement toward the goal of attaining it, and not simply in assessment of the discrepancy (Kumashiro et al. 2007). The assessment approach involves a primary focus on evaluation and critique, and when one is the target of the sculpting process (you have a partner who simply assesses your discrepancy to your ideal self), this leads one to think the other is not supportive. When one is the sculptor, the assessment approach leads one to be less affirming, disapproving of the goals, less involved in the sculptee's pursuits. Mere assessment is debilitating to the sculpting process. Locomotion involves a

¹²Like many of the concepts reviewed in this book, the socially constructed self has a long theoretical history going back to Aristotle's emphasis on friendship, social science formulated by James (1890), Cooley (1902), and Mead (1934), and continental thought from Kierkegaard (1846/1992) to Jaspers (1919) to Gadamer (1975).

focus on affirmation and moving toward the goal of the sculptee. This effect may well be culture-dependent (Markus & Kitayama, 2010).

Research on the Michelangelo phenomenon has only recently been extended to influence among friends and between supervisor–supervisee work pairs (Bagci et al., 2016) and it also appears to hold in these relationships.¹³ Thus, the Michelangelo phenomenon may provide a window into one way that social influence in communities, organizations, and culture may work. The influence of the local unit in organizations comes through close, interdependent interactions with coworkers. The influence of culture comes from socialization experiences with proximate others too. The Michelangelo phenomenon may play a central role in much of the influence that Markus and Kitayama (2010) propose in the “daily situations and practices” aspect of their model of cultural influence on the self. There are, of course other levels to the model of culture and self (language, media, education, pervasive ideas, etc.), but this cooperative sculpting process may play a central role in how we become (or avoid or fail to become) moral.

3.5 Influencing the Moral Ecology

The moral ecology metaphor encourages us to think about cultural influence as reciprocal. We often have the freedom to choose our influencers, to volunteer, to join organizations, to move cultures, to protest our treatment, to encourage moral action in others, to try to change our surroundings or to leave them behind. In short, we often have the freedom to influence both the small- and (less often) large-scale moral ecology that surrounds us.

3.5.1 Organizational Ethics As a Design Problem

Epley and Tannenbaum (2017) propose we construe ethical action in organizations as a design problem. How do we design organizations so that they promote and produce ethical action? They review a large literature in organizational psychology (some of which is included in Section 3.3) and then propose that organizations can construct hiring, compensation, reputation management, and operations processes so that they support ethical action. They list three psychological levels on which this ethical design can occur: attention, construal, and motivation.

3.5.1.1 Attention: Are Ethics Top of Mind? In organizations attention is often focused on goals of efficiency or profit. But organizational processes can be designed to make ethical concern and evaluation a part of routine decision processes. They offer the hiring interview protocol that companies like Johnson & Johnson use. The interview is structured in part to begin the socialization process of new employees into corporate culture. It includes numerous questions focusing on how prospective employees

¹³The process may also be similar to some kinds of parental influence (Patrick & Gibbs, 2012).

might support the company value of prioritizing the needs of the people it serves. This focuses attention on the values that are held in that organization.

3.5.1.2 Construal: Are People Asking “is It Right?” The procedures that a company follows are often simply implemented without ever being evaluated. But do employees ever ask, “Is this the right thing to do?” Compensation policies that tie executive performance to clear measures of values can help connect personal decisions made by executives to values the organization supports. Southwest Airlines, for instance, has an “every employee matters” slogan that is incorporated into the compensation packages of executives by tying compensation to measures of voluntary turnover in that executive’s division.

3.5.1.3 Motivation: Are You Using Pro-social Goals? Most employees are interested in doing (or at least seeing) some good in the work they do. Organizations can leverage this already existing desire and align it with values the organization supports. Virgin Atlantic rewarded its pilots for increasing fuel efficiency by donating money to charities of the pilot’s choice. This change in policy significantly increased job satisfaction.

It seems clear there are ways for organizations to influence the ethical behavior of their employees and executives. These approaches are fraught with complexity (e.g. the paradox of control) but they can produce significant organizational change and support moral action within organizational cultures.

3.5.2 Moral Exemplars and Social Reformers

The pioneering work of Colby and Damon (1992) in moral exemplars shows great variety of goals among their exemplars, even though their sample was primarily limited to individuals doing social service. Among their exemplars were people who campaigned for justice in civil rights, for health care, or for equal economic opportunity. Huff and Barnard (2009) found a similar distinction among cases in the computing field. The computer scientists, software engineers, and policy advocates they interviewed included a significant number of reformers campaigning for opensource software, privacy policy, inclusion of women in the field, safety standards in software, and other goals. In both studies, those working for social change often expressed frustration at the opposition they faced, and the difficulties involved in producing social change. Working for social change is difficult, and these exemplars leveraged social support and their own optimism to maintain their motivation.

Hirschman’s (1970) *Exit, Voice and Loyalty* provides a structure for the choices facing those who work in firms, organizations, and government as these begin to conflict with the values and principles of members. The framework (with the addition of a “neglect” option) has been used to explain a wide variety of choices for and against social change including the decision to blow the whistle in government and business (Uys, 2022), organizational change, and human resources (Allen, 2014). These choices resonate most clearly with the case of the 2020 presidential election in the United States that we outlined in the Introduction. Different actors in the Republican Party chose differently, some opting for voice (as did the two in our case),

some opting for exit (as did some of their colleagues who resigned), and some opting for loyalty to the president in his last days in office.

3.5.3 Moral Action as Critique and as Design

Leadership and social change work are classic examples of how the moral ecology is shaped, but there are many other informal approaches. The three processes we reviewed in Section 3.3.2 – rationalization, institutionalization, and socialization – are also modes of influence to change the moral ecology for the better or worse. One can use cognitive and motivational strategies to make stakeholders more (or less) visible, risks more (or less) considered, and choices more (or less) conscious (Epley & Tannenbaum, 2017; Messick, 1996). One can institutionalize these strategies in offices of ethics and compliance, or in the process models used in the development of products. And one can socialize individuals into these processes.

When organizations or culture do not meet one's moral standards, one can attempt social change (choosing voice over exit or loyalty). The work on volunteerism, social reform, and, as a last hope, courageous resistance provides examples of this strategy that are widely spread across cultures (see the Introduction for examples). Much of the work of moral exemplars in social service (Colby & Damon, 1992) and computing (Huff & Barnard, 2009) consists of trying to change moral ecologies, often in face of stiff resistance. In addition, when others offend the moral standards of a community, individuals in that community can take it upon themselves to defend the standards by punishing the offender, often at cost to themselves (Bowles et al., 2012; Boyd, Gintis, & Bowles, 2010; Boyd et al., 2003). Even children engage in vocal protest against injustice in social relations, ever more effectively as they develop skills such as perspective taking (Killen & Dahl, 2021).

3.5.4 Leadership

One place where we might expect this design to happen is in leadership in organizations. In agreement with the emphasis here on ecology, models of leadership have moved from classical great-man and matching approaches to an understanding of the complexity of leadership in organizations as a “dynamic cocreational process among leaders, followers, and environments” (Thoroughgood et al., 2016, p. 627). Leadership is now recognized as being shared and distributed, even in hierarchical and control-oriented organizations (Avolio, Walumbwa, & Weber, 2009; Hogg, 2001). And the process by which leaders influence others has been expanded to include emotion, value, and self-concept variables in addition to classical control models (Avolio et al., 2009). For instance, work by Vianello and Haidt (2010) shows that one way that leaders influence others is through eliciting the moral emotion of *elevation* in response to the leader's display of interpersonal fairness and self-sacrifice. Bad leadership can produce negative moral emotions (e.g. anger, disgust) and these emotions can be contagious, spreading to other workers without the need for direct negative experience (Dasborough et al., 2009). In this respect, moral potency has been suggested as a central measure of a leader's effectiveness in inducing change, and is the combined effect of courage, efficacy,

and ownership (Hannah & Avolio, 2010). However, it concentrates narrowly on characteristics of the leader.

An important lesson emerges from the evolution of leadership theories. One must beware of romanticizing a role as primarily constituted by the excellent or expert characteristics of the person inhabiting that role. There has been a trenchant critique of this romanticizing of the leader in leadership research (Collinson, Jones, & Grint, 2018; Ford, Harding, & Gilmore, 2022; Meindl, Ehrlich, & Dukerich, 1985) and a concomitant focus on the complexity of the social interactions among leaders, followers, and the environment. Hogg (2001), for example, uses social identity theory to construct a model of leadership that is a function of the social system rather than the leader. Shared social identity and an in-group-based influence process identify some members as prototypical leaders thereby producing in-group status differences. While both leaders and followers negotiate over the values, purposes, and plans of the group, there is an attributional process that misattributes all, or most of, the influence to the leader (Hogg, 2001).

The literature on moral exemplars and moral expertise needs to take note of this critique, and this chapter is the right place to do so. We know very little of how moral exemplars interact with their collaborators, followers, and the surrounding moral ecology. Colby and Damon's (1992) classic study of exemplars looks carefully at their exemplars' interactions with others as a part of their developmental trajectory; but it still treats the exemplariness of their action as mostly originating from their exemplary character rather than from the "dynamic cocreational process" between exemplars and their moral ecology. One can find the beginning of this critique in some places in the exemplar literature (Frimer, 2018; King, Mueller, & Furrow, 2013).

With this section on influencing the moral ecology, we come full circle in the cycle of mutual constitution of individuals and culture (Hamedani & Markus, 2019; Markus & Kitayama, 2010). This is a process constantly in flux, at a variety of levels of influence and resistance, and we can only gesture to its complexity here.

3.6 Discussion

This review of the ways that moral ecology can influence moral action, and be influenced by it, is admittedly spotty. The point is not to make a comprehensive review of moral psychology as influenced by friendship, families, groups, organizations, and culture.¹⁴ It is more to show the wide variety of ways that moral action can be supported or led astray by the social surround, and the ways that the social environment can be influenced so it is more supportive of moral action. If the review provides windows into the processes of moral ecology and some idea of its variety and importance, then it has been successful. Still, there are some tentative conclusions we can make from the work reviewed so far.

¹⁴This would be a book in itself and we know of no volume attempting this.

3.6.1 Conclusion

The range of moral ecologies is wide, but they share characteristics across their variety. The extent to which these characteristics are shared across all levels is a matter for empirical investigation, so these conclusions should be taken as tentative. Some conclusions (e.g. Section 3.6.1.1) share an evidence base from other chapters and can be thought of as more established:

1. *Moral psychology is also the psychology of evil.*

Many of the moral ecology processes we have covered have a complex relationship to the content of the moral action they support or hinder. Designing institutional processes to support particular valued employee and executive actions might be a way to support either corruption or ethical action. The limits of the Michelangelo phenomenon have yet to be explored but it seems reasonable that it would work for one partner to help another to obtain an ideal self that might harm the self (e.g. anorexia) or others (e.g. terrorism). We may find some processes in moral psychology that are more likely to result in immoral rather than moral action. For instance, the moral disengagement strategy of minimizing harm is pervasive in justifications of aggression. Still, surgeons and doctors may need to psychologically minimize the harm they do to allow themselves permission to do surgery (Collins & Pinch, 2005). And soldiers in combat need to objectivize the enemy to do their job (Grossman, 1996). In this way, moral psychology is likely to be immoral psychology, and the psychology of moral exemplars also becomes the psychology of evil.

2. *Moral ecologies can support domain limits in moral judgment and action, making integrity across situations difficult.*

A specific moral ecology can provide domain-limited morally relevant information (Is this a moral issue? Are these people I should care about?). It can also prime particular moral emotions and skills and thereby make certain kinds of (im)moral action more relevant and/or likely. Cultural institutions, values, activities, organizational procedures, domains of concern in volunteer organizations, and the moral support of close others can define particular issues as important or irrelevant to the moral self. It can even make some individuals subject to compassion and others subject to attack (Hart, 2005). Some moral principles (e.g. group loyalty, Haidt & Joseph, 2004) have this domain limitation as an integral part of their structure. This may be a more general problem about the limitation of human resources: focus on anything requires one to remove focus from something else.

3. *There is a structure to the moral differences both across and within cultures.*

Our catalog of moral differences within and across cultures is already quite long even in this short review. But what is often missed is that these cultural differences have a structure to them that allows one to see a larger pattern to the agreement and disagreement within and across cultures. There are still multiple ways to describe the dimensions in these differences, and it may take some time to determine the best empirical solutions to a description. But there is agreement that there is indeed a structure and that cultures and individuals do not simply randomly choose to value some things (e.g. the color blue) and not others. This makes the empirical case for radical “anything goes” cultural relativism less persuasive.

A more likely description is that there is some finite set of values or dimensions of valuing that are framed by evolution or the constraints of social interaction and that individuals and societies make choices within this limited space.

4. *Moral diversity within and across cultures is both a problem and a solution.*

Moral diversity and disagreement can be found within every culture and across cultures.¹⁵ This has usually been approached as a problem. From a philosophical standpoint, it looks too much like support for unthinking relativism. From a political and religious standpoint, it seems the nexus of a great deal of societal conflict and uncertainty. However, Lawrence Kohlberg's work and that of many others reviewed in this and other chapters (Chapters 7, 8, and 9) make the clear point that this diversity serves as a resource for cultural critique. The shared structure and values within and across cultures make conversations possible across these value divides. They serve as resources for arguments that appeal to common experience, value, and intuition in the face of disagreement. This diversity may indeed be seen as similar in importance to genetic diversity in evolutionary processes: without the existing diversity, change would be impossible.

5. *One can approach morality as a design issue.*

Because (im)moral decisions are imbedded in the structures of relationships, families, organizations, and cultures, one can think of ways to change those relationships so that they are more likely to result in moral behavior. This approach connects current research on moral action to ancient traditions of moral thought from Greece, India, and China. Philosophers have long thought about how to structure society. This thought tradition makes it clear that the "design problem" approach has no easy solution, nor does it guarantee moral outcomes. But current work in moral psychology helps us discern some structure to the design problem that is based in the psychological capacities and motivations of people.

6. *Recognition of diversity within and across cultures does not require radical relativism.*

The existence of cultural difference is indisputable, as are the patterns of those differences across cultures: cultures agree on which principles they disagree about, and these disagreements have patterns (e.g. values that tend to be endorsed or sanctioned in cultures cluster together, see Schwartz, 2010). These differences may have their origins in evolutionary patterns required for successful group living (Tomasello & Vaish, 2012) and so to that extent they are objective in terms of embodying what is needed for flourishing for the species (Flanagan, 1991; Flanagan et al., 2008). Thus, recognition of these patterns of cultural difference does not require acceptance of cultural relativism or an "anything goes as long as a group endorses it" approach (Flanagan et al., 2008; Taylor, 1982). It is at least compatible with a suitably complex pluralism, a recognition that there are in the human condition multiple foundational values and that different cultures come to different solutions about the appropriate balance among them. A pluralistic approach like this may help in guarding against the sort of moral objectivism (Goodwin & Darley, 2008, 2009, 2010, 2012) that associates a person's culture's judgments and actions unequivocally with the good, and demonizes those who disagree (Kekes, 1993).

¹⁵My thanks to Jessica Salvatore, one of the reviewers of this text, for this insight.

7. *Humans are not passive recipients of influence in culture, but active co-creators of it.* In almost every section of this chapter we saw humans being influenced by culture, organization, leaders, friends, etc. However, we also saw multiple examples of people changing organizations, influencing each other, and taking action to support other people taking moral action. We provide links to some models of how this happens (e.g. Markus & Kitayama, 2010) but more work is needed.

3.6.2 Application

The story of Nathaniel Borenstein's decision presented in the Introduction is a story of conflicting and overlapping moral ecologies. In it we have a happy incident in which we see multiple moral ecologies influencing an actor to disregard his principles and thereby achieve a moral outcome. Borenstein is a computer scientist and pacifist with a long history in the peace movement. His perception of the moral ecology of the military was that it was indifferent if not hostile to his concerns about the safety of nuclear weapons systems. But that very system recruited him to help it critique its own performance on safety. He was motivated to help the military in its self-critique despite his firm personal value commitment to not consulting for the military. As he began his collaboration, he discovered that the value conflict was more that some actors in the private consulting sector were unconcerned with safety if they could sell a contract. His action in altering military perception and policy about certain computer systems is an example of committed individuals taking moral action to change moral ecologies. The NATO lieutenant commander who persistently recruited Borenstein was a pivotal player, as was Borenstein himself. Each had different psychologies, goals, and moral influences, but they shared significant overlap that helped to produce conversation and change. The story is also an example of one moral ecology (NATO) using a representative of another moral ecology (the peace movement) to influence a third moral ecology (the military consulting industry). That these very different moral ecologies had something to say to each other is evidence of the cross-cultural similarities in the structure of values and an example of how differences are negotiated.

3.6.3 Open Questions

In this chapter, we have focused on the mutual influence that people and their surrounding moral ecologies exert on each other. This has led us to consider the complexities of how the various levels of moral ecology interact with each other. Most of these complexities are just beginning to be investigated.

1. *Our knowledge of how groups take moral action is fragmented. Addressing this may move us from thinking of virtuous individuals to virtuous ecologies, or the interaction of these, as the source of effective virtue.*

Work on morality in groups is often isolated from work in moral psychology. In this chapter, we have mostly talked about moral ecology from the perspective of the individual moral actor and his or her role in the moral ecology. But it is possible to move to a group level of analysis. We need better conceptualizations of the group processes involved in moral awareness, judgment, motivation, and action that can help drive research at this level. Some of the work mentioned earlier in leadership

(Section 3.5.4) might be a source from which to begin theorizing. We have specifically called for work in moral exemplars to start to account for this systemic level of exemplary action. With moral exemplars, one might focus on continued exemplary action (rather than just the single actor) and ask what systems support and hinder it. This might move the conceptual balance of moral character to seeing the “dynamic cocreational process among leaders, followers, and environments” (Thoroughgood et al., 2016, p. 627) as a source of effective virtue.

2. *How do moral actors actively grapple with moral diversity?*

Clearly, cultures, organizations, and even friendships are not homogenous. It is at least clear that the reaction to moral diversity at all these levels is often negative: we judge those who have moral disagreements with us harshly. We need to know more about how people think about, support, or tolerate the diversity of moral judgments and goals in a moral ecology. How do people handle moral tradeoffs when diverse values are in play? How do choice processes lead to different moral goals at these different levels? How do individuals lobby to turn optional moral goals into mandatory goals? There are surely even more questions than these.

3. *How do the different levels of a moral ecology interact with each other?*

Markus and Kitayama (2010) have given a beginning conceptualization of how self and culture mutually constitute each other, and their model involves interactions at all the levels we have reviewed in this chapter. They include internal representations of the moral ecology, for instance, how daily routines (e.g. hygiene) should be structured, or pervasive moral ideas such as how dependent or interdependent people should be with others. However, their model does not yet seem specific enough to drive programs of research in morality (e.g. the word moral only occurs twice in the article).

We also do not know enough about the consistency of the structure of value differences across different levels of a moral ecology: Are self-transcendence and self-enhancement always viewed as in conflict in all cultures at all levels of a moral ecology? Nor do we know much about how the larger moral ecology structures the relationships of its constituent units: Why are there such large differences in the frequency of volunteerism across cultures?

4. *How does the active construction of identity interact with the various levels of moral ecology?*

We know from Chapter 5 that how (and whether) people think of themselves in moral terms is of crucial importance. And we know that one cannot be a self by oneself. What does this mean for how a moral ecology helps to constitute the moral identities of those within it? We know there are significant moral differences within moral ecologies. How are moral identities related to those differences, and how do they construe the differences in the ecology?

3.7 Further Readings

These suggested reading are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Ashforth and Anand (2003). “The normalization of corruption in organizations.” A literature review and a model of how organizations shape immoral behavior.

- Epley and Tannenbaum (2017). “Treating ethics as a design problem.” A proposal that ethical issues in organizations can be treated as a design issue for them, making the point that an organization can shape the ethical behavior of its members.
- Heine (2010). “Cultural psychology.” A readable overview of the work on psychological differences across culture.
- Markus and Kitayama (2010). “Cultures and selves: A cycle of mutual constitution.” This proposes a multilayered model of how culture influences the self and how the self influences culture.
- Nisbett and Ross (1980). *Human Inference: Strategies and Shortcoming of Social Judgment*. The classic early statement of the primacy of situational influence on judgment and decision-making.
- Sagiv and Schwartz (2022). “Personal values across cultures.” A comprehensive review of current work on values, focusing on Schwartz’s theory of values. It explores cross-cultural variation and similarity and the interlinking of cognitive representations, emotions, and motivational goals.

References

- Allen, M. M. C. (2014). Hirschman and Voice. In A. Wilkinson, J. Donaghey, T. Dundon, & R. Freeman (Eds.), *The Handbook of Research on Employee Voice* (pp. 36–51). New York: Edward Elgar.
- Anon. (2100 BCE/1989). *The Epic of Gilgamesh* (M. G. Kovacs, Trans.). Stanford, CA: Stanford University Press.
- Aristotle. (340 BCE/1941). Nicomachean Ethics (W. D. Ross, Trans.). In R. McKeon (Ed.), *The Basic Works of Aristotle* (pp. 927–1112). New York: Random House.
- Ashforth, B., & Anand, V. (2003). The normalization of corruption in organizations. *Research in Organizational Behavior*, 25, 1–52. doi:10.1016/s0191-3085(03)25001-2
- Association of Certified Fraud Examiners. (2018). Report to the nations: 2018 study on occupational fraud and abuse. Retrieved from <https://s3-us-west-2.amazonaws.com/acfepublic/2018-report-to-the-nations.pdf>
- Avolio, B. J., Walumbwa, F. O., & Weber, T. J. (2009). Leadership: Current theories, research, and future directions. *Annual Review of Psychology*, 60, 421–449.
- Bagci, S. C., Kumashiro, M., Rutland, A. et al. (2016). Cross-ethnic friendships, psychological well-being, and academic outcomes: Study of South Asian and White children in the UK. *European Journal of Developmental Psychology*, 14(2), 190–205. doi:10.1080/17405629.2016.1185008
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, 31(2), 101–119.
- Bardi, A., & Schwartz, S. H. (2003). Values and behavior: strength and structure of relations. *Personality and Social Psychology Bulletin*, 29(10), 1207–1220. doi:10.1177/0146167203254602
- Bellah, R. N. (2011). *Religion in Human Evolution: From the Paleolithic to the Axial Age*. Cambridge, MA: Belknap.
- Bloom, P. (2012). Religion, morality, evolution. *Annual Review of Psychology*, 63, 179–199. doi:10.1146/annurev-psych-120710-100334
- Bowles, S., Boyd, R., Matthew, S. et al. (2012). The punishment that sustains cooperation is often coordinated and costly. *Behavioral and Brain Sciences*, 35(1), 20–21.

- Bowman, N. A., Kitayama, S., & Nisbett, R. E. (2009). Social class differences in self, attribution and attention: Socially expansive individualism of middle-class Americans. *Personality and Social Psychology Bulletin*, 35(7), 880–893.
- Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328(5978), 617–620. doi:10.1126/science.1183665
- Boyd, R., Gintis, H., Bowles, S. et al. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6), 3531–3535. doi:10.1073/pnas.0630443100
- Braithwaite, J. (1989). Criminological theory and organizational crime. *Justice Quarterly*, 6(3), 333–358.
- Brief, A. P., Buttram, R. T., & Dukerich, J. M. (2001). Collective corruption in the corporate world: Toward a process model. In M. E. Turner (Ed.), *Groups at Work: Theory and Research* (pp. 471–499). Mahwah, NJ: Lawrence Erlbaum.
- Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, 81(4), 697–710.
- Cohen, A. B., Rozin, P., & Keltner, D. (2004). Different religions, different emotions. *Behavioral and Brain Sciences*, 27(6), 734–735.
- Cohen, A. B., Siegel, J. I., & Rozin, P. (2003). Faith versus practice: different bases for religiosity judgments by Jews and Protestants. *European Journal of Social Psychology*, 33(2), 287–295. doi:10.1002/ejsp.148
- Colby, A., & Damon, W. (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. New York: Free Press.
- Collins, H., & Pinch, T. (2005). *Dr. Golem: How to Think About Medicine*. Chicago, IL: University of Chicago Press.
- Collins, H., & Pinch, T. (2006). *Dr. Golem: How to Think About Medicine*. Chicago, IL: University of Chicago.
- Collinson, D., Jones, O. S. c., & Grint, K. (2018). “No more heroes”: Critical perspectives on leadership romanticism. *Organization Studies*, 39(11), 1625–1647. doi:10.1177/0170840617727784
- Cooley, C. H. (1902). *Human Nature and the Social Order*. New York: Charles Scribner’s Sons.
- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, 60(1), 47–69. doi:10.1086/701478
- Darley, J. M. (1994). Gaming, gundecking, body counts, and the loss of three British cruisers at the battle of outland: The complex moral consequences of performance measurement systems in military settings. Unpublished speech to Air Force Academy. Princeton University.
- Dasborough, M. T., Ashkanasy, N. M., Tee, E. Y. J. et al. (2009). What goes around comes around: How meso-level negative emotional contagion can ultimately determine organizational attitudes toward leaders. *The Leadership Quarterly*, 20(4), 571–585. doi:10.1016/j.leaqua.2009.04.009
- Davis, M. (Ed.) (1998). *Thinking Like an Engineer: Studies in the Ethics of a Profession*. New York: Oxford University Press.
- Denison, D. R. (1996). What is the difference between organizational culture and organizational climate? A native’s point of view on a decade of paradigm wars. *Academy of Management Review*, 619–654.
- Dobbin, F., & Kalev, A. (2016). Why diversity programs fail. *Harvard Business Review*, (July–August), 1–23.
- Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

- Doris, J. M., & Stich, S. P. (2005). As a matter of fact: Empirical perspectives on ethics. In F. Jackson & M. Smith (Eds.), *The Oxford Handbook of Contemporary Philosophy*, 114–152.
- Downey, G. L., Lucena, J. C., & Mitcham, C. (2007). Engineering ethics and identity: Emerging initiatives in comparative perspective. *Science and Engineering Ethics*, 13(4), 463–487.
- Drigotas, S. M., Rusbult, C. E., Wieselquist, J. et al. (1999). Close partner as sculptor of the ideal self: Behavioral affirmation and the Michelangelo phenomenon. *Journal of Personality and Social Psychology*, 77(2), 293–323.
- Durkheim, É. (1925/1973). *Moral Education: A Study in the Theory and Application of the Sociology of Education*. New York: Free Press.
- Eberhardt, N. (2014). Piaget and Durkheim: Competing paradigms in the anthropology of morality. *Anthropological theory*, 14(3), 301–316. doi:10.1177/1463499614534552
- Epley, N., & Tannenbaum, D. (2017). Treating ethics as a design problem. *Behavioral Science & Policy*, 3(2), 73–84.
- Flanagan, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press.
- Flanagan, O., Sarkissian, H., & Wong, D. (2008). Naturalizing Ethics. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Evolution of Morality: Adaptations and Innateness* (vol. 1, pp. 1–26). Cambridge, MA: MIT Press.
- Ford, J., Harding, N., & Gilmore, S. (2022). RE/searching leadership: A critique in two agonies and nine fits. *Human Relations*, 0(0). doi.org/10.1177/00187267221079167
- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330–347.
- Frimer, J. A. (2018). Moral heroes are puppets. In K. Gray & J. Graham (Eds.), *The Atlas of Moral Psychology* (pp. 252–257). New York: Guilford Press.
- Gadamer, H.-G. (1975). *Wahrheit und Methode. Grundzüge einer philosophischen Hermenutik*. Tübingen: J. C. B. Mohr.
- Gehman, J., Treviño, L. K., & Garud, R. (2013). Values work: A process study of the emergence and performance of organizational values practices. *Academy of Management Journal*, 56(1), 84–112.
- Gibbs, J. C. (1977). Kohlberg's stages of moral judgment: A constructive critique. *Harvard Educational Review*, 47, 43–61.
- Gibbs, J. C., Basinger, K. S., Grime, R. L. et al. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review*, 27(4), 443–500. doi:10.1016/j.dr.2007.04.001
- Ginges, J., Hansen, I., & Norenzayan, A. (2009). Religion and support for suicide attacks. *Psychological Science*, 20(2), 224–230. doi:10.1111/j.1467-9280.2009.02270.x
- Goodwin, G. P., & Darley, J. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition*, 106(3), 1339–1366. doi:10.1016/j.cognition.2007.06.007
- Goodwin, G. P., & Darley, J. M. (2009). The perceived objectivity of ethical beliefs: Psychological findings and implications for public policy. *Review of Philosophy and Psychology*, 1(2), 161–188. doi:10.1007/s13164-009-0013-4
- Goodwin, G. P., & Darley, J. M. (2010). The perceived objectivity of ethical beliefs: Psychological findings and implications for public policy. *Review of Philosophy and Psychology*, 1, 161–188.
- Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology*, 48(1), 250–256. doi:10.1016/j.jesp.2011.08.006
- Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. R. Shaver & M. Mikulincer (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil* (pp. 11–31). New York: American Psychological Association Books.

- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029–1046.
- Gregg, G. S. (2005). *The Middle East: A Cultural Psychology*. New York: Oxford University Press.
- Grimpe, B., Stahl, B., Ten Holter, C. et al. (2020). From collaborative to institutional reflexivity: calibrating responsibility in the funding process. *Science and Public Policy*, 47(5), 720–732. doi:10.1093/scipol/scaa038
- Grossman, D. (1996). *On Killing: The Psychological Cost of Learning to Kill in War and Society*. New York: Back Bay Books.
- Haeckel, E. H. P. A. (1866). *Generelle Morphologie der Organismen: Zweiter Band* (vol. 2). Berlin: Georg Reimer.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J. (2010). Moral psychology must not be based on faith and hope: Commentary on Narvaez (2010). *Perspectives on Psychological Science*, 5(2), 182.
- Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Pantheon.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.
- Haidt, J., & Kesebir, S. (2010). Morality. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (5th ed., pp. 787–832). Hoboken, NJ: Wiley.
- Hamedani, M. Y. G., & Markus, H. R. (2019). Understanding culture clashes and catalyzing change: A culture cycle approach. *Frontiers in Psychology* 10, 700. doi:10.3389/fpsyg.2019.00700
- Hannah, S., T., & Avolio, B. J. (2010). Moral potency: Building the capacity for character-based leadership. *Consulting Psychology Journal*, 62(4), 291–310.
- Hart, D. (2005). Adding identity to the moral domain. *Human Development*, 48, 257–261.
- Hart, D., Atkins, R., & Southerland, N. (2006). Community service and moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 633–656). Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers.
- Hart, D., & Carlo, G. (2005). Moral development in adolescence. *Journal of Research on Adolescence. Special Issue: Moral Development*, 15(3), 223–233. doi:10.1111/j.1532-7795.2005.00094.x
- Hart, D., & Fegley, S. (1995). Prosocial behavior and caring in adolescence: Relations to self-understanding and social judgment. *Child Development*, 66, 1346–1359.
- Hart, D., & Matsuba, M. K. (2007). The development of pride and moral life. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The Self-conscious Emotions: Theory and Research* (pp. 114–133). New York: Guilford Press.
- Hart, D., & Matsuba, M. K. (2009). Urban neighborhoods as contexts for moral identity development. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 214–231). New York: Cambridge University Press.
- Healy, P., & Serafeim, G. (2012). Causes and consequences of firm disclosures of anticorruption efforts. Harvard Business School working paper 12-077.
- Heine, S. J. (2010). Cultural psychology. In S. T. Fiske, D. T. Gilbert, & L. Gardner (Eds.), *Handbook of Social Psychology* (5th ed., pp. 1423–1464). New York: John Wiley.
- Helton-Fauth, W., Gaddis, B., Scott, G. et al. (2003). A new approach to assessing ethical conduct in scientific work. *Accountability in Research*, 10(4), 205–228.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. doi:10.1017/s0140525x0999152x.
- Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Thousand Oaks, CA: Sage.
- Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review* 5(3), 184–200. doi:10.1207/S15327957PSPR0503_1
- Huff, C., & Barnard, L. K. (2009). Good computing: moral exemplars in the computing profession. *IEEE Technology and Society Magazine* (Fall), 47–54.
- Huff, C., Barnard, L. K., & Frey, W. (2008a). Good computing: A pedagogically focused model of virtue in the practice of computing (part 1). *Journal of Information, Communication and Ethics in Society*, 6(3), 246–278.
- Huff, C., Barnard, L. K., & Frey, W. (2008b). Good Computing: A pedagogically focused model of virtue in the practice of computing (part 2). *Journal of Information, Communication & Ethics in Society*, 6, 284–316.
- Inglehart, R., & Welzel, C. (2005). *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.
- James, W. (1890). *The Principles of Psychology*. New York: Holt.
- Jaspers, K. (1919). *Psychologie der Weltanschauungen*. Berlin: Springer.
- Jones, R. P. (2016). *The End of White Christian America*. New York: Simon & Schuster.
- Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *The Academy of Management Review*, 16(2), 366–395.
- Kekes, J. (1993). *The Morality of Pluralism*. Princeton, NJ: Princeton University Press.
- Kierkegaard, S. (1846/1992). *Concluding Unscientific Postscript to Philosophical Fragments: A Mimetic-Pathetic-Dialectic Compilation, an Existential Plea, by Johannes Climacus* (H. V. Hong & E. H. Hong, Trans.). Princeton NJ: Princeton University Press.
- Killen, M., & Dahl, A. (2021). Moral reasoning enables developmental and social change. *Perspectives on Psychological Science*, 16(6), 1209–1225. doi:10.1177/1745691620964076
- King, P. E., Mueller, R. A., & Furrow, J. (2013). Cultural and contextual issues in exemplar research. *New Directions for Child Adolescent Development*, 2013(142), 41–58. doi:10.1002/cad.20048
- Kinnier, R. T., Kernes, J. L., & Dautheribes, T. M. (2000). A short list of universal moral values. *Counseling and Values*, 45(1), 4–16. doi:10.1002/j.2161-007X.2000.tb00178.x
- Kitayama, S., Ishii, K., Imada, T. et al. (2006). Voluntary settlement and the spirit of independence: Evidence from Japan's "northern frontier." *Journal of Personality and Social Psychology*, 91(3), 369–384. doi:http://dx.doi.org/10.1037/0022-3514.91.3.369
- Kohlberg, L. (1963). The development of children's orientations toward a moral order: I. Sequence in the development of moral thought. *Vita Humana*, 6(1–2), 11–33.
- Kohlberg, L., & Mayer, R. (1972). Development as the aim of education. *Harvard Educational Review*, 42(4), 449–496.
- Kumashiro, M., Rusbult, C. E., Finkenauer, C. et al. (2007). To think or to do: The impact of assessment and locomotion orientation on the Michelangelo phenomenon. *Journal of Social and Personal Relationships*, 24(4), 591–611. doi:10.1177/0265407507079261
- Leidner, D. E. (2006). Review: A review of culture in information systems research: Toward a theory of information technology cultural conflict. *MIS Quarterly*, 30(2), 357–399.
- Leveson, N. G. (1995). *Safeware: System Safety and Computers*. New York: Addison-Wesley.
- Levine, S., Rottman, J., Davis, T. et al. (2021). Religious affiliation and conceptions of the moral domain. *Social Cognition*, 39(1), 139–165. doi:10.1521/soco.2021.39.1.139
- Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, 5(4), 420–430. doi:10.1177/1745691610375557
- Marty, M., & Appleby, R. S. (Eds.). (1991). *Fundamentalisms Observed. The Fundamentalism Project*, vol. 1. Chicago, IL: University of Chicago Press.
- Matsuba, M. K., Hart, D., & Atkins, R. (2007). Psychological and social-structural influences on commitment to volunteering. *Journal of Research in Personality*, 41(4), 889–907.

- Matsuba, M. K., & Walker, L. J. (2004). Extraordinary moral commitment: Young adults involved in social organizations. *Journal of Personality*, 72(2), 413–436.
- Meindl, J. R., Ehrlich, S. B., & Dukerich, J. M. (1985). The romance of leadership. *Administrative Science Quarterly*, 30(1), 78–102. doi:10.2307/2392813
- Messick, D. M. B., & Max H. (1996). Ethical leadership and the psychology of decision making. *Research in Ethical Issues in Organizations*, 3, 213–238.
- Milgram, S. (1963). Behavioral Study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371–378.
- Milgram, S. (1974). *Obedience to Authority: An Experimental View*. New York: Harper and Row.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.
- Moore, D. A., Tetlock, P. E., Tanlu, L. et al. (2006). Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review*, 31(1), 10–29.
- Muthukrishna, M., Bell, A. V., Henrich, J. et al. (2020). Beyond western, educated, industrial, rich, and democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, 31(6), 678–701. doi:10.1177/0956797620916782
- Nisbett, R. E., & Cohen, D. (1996). *Culture of Honor: The Psychology of Violence in the South*. Boulder, CO: Westview Press.
- Nisbett, R. E., & Ross, L. (1980). *Human Inference: Strategies and Shortcoming of Social Judgment*. Mahwah, NJ: Prentice Hall.
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32–48.
- Oliner, S. P., & Oliner, P. M. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press
- Paluck, E. L., Porat, R., Clark, C. S. et al. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72, 533–560.
- Park, R. E. (1952). *Human Communities: The City and Human Ecology*. Glencoe, IL: Free Press.
- Patrick, R. B., & Gibbs, J. C. (2012). Inductive discipline, parental expression of disappointed expectations, and moral identity in adolescence. *Journal of Youth and Adolescence*, 41(8), 973–983. doi:10.1007/s10964-011-9698-7
- Power, F. C. (2004). The moral self in community. In D. K. Lapsley & D. Narvaez (Eds.), *Moral Development, Self and Identity* (pp. 47–64). Mahwah, NJ: Lawrence Erlbaum.
- Rauthmann, J., & Sherman, R. A. (2020). The situation of situation research: Knowns and unknowns. *Current Directions in Psychological Science*, 29(5), 473–480. doi:10.1177/0963721420925546
- Reimer, K. S., DeWitt Goudelock, B. M., & Walker, L. J. (2009). Developing conceptions of moral maturity: Traits and identity in adolescent personality. *The Journal of Positive Psychology*, 4(5), 372–388. doi:10.1080/17439760902992431
- Richards, P. S., & Davison, M. L. (1992). Religious bias in moral development research: A psychometric investigation. *Journal for the Scientific Study of Religion*, 31(4), 467–485.
- Roccas, S., Sagiv, L., Schwartz, S. H. et al. (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin*, 28(6), 789–801.
- Rusbult, C. E., Finkel, E. J., & Kumashiro, M. (2009). The Michelangelo phenomenon. *Current Directions in Psychological Science*, 18, 305–309.
- Sagiv, L., & Schwartz, S. H. (2022). Personal values across cultures. *Annual Review of Psychology*, 73, 517–546. doi:10.1146/annurev-psych-020821-125100
- Schroeder, D. N. (1992). Aristotle on the good of virtue-friendship. *History of Political Thought*, 13(2), 203–218.

- Schwartz, S. H. (2006). Basic human values: Theory, measurement and applications. *Revue française de sociologie*, 47(4), 929–968.
- Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives Emotions, and Behavior: The Better Angels of Our Nature* (pp. 221–241). Washington, DC: American Psychological Association.
- Shilton, K. (2018). Values and ethics in human–computer interaction. *Foundations and Trends in Human–Computer Interaction*, 12(2), 107–171. doi:10.1561/11000000073
- Shweder, R. A., Much, N. C., Mahapatra, M. et al. (1997). The “big three” of morality (autonomy, community, and divinity) and the “big three” explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and Health* (pp. 119–169). New York: Routledge.
- Snow, N. E. (2010). *Virtue as Social Intelligence: An Empirically Grounded Theory*. New York: Routledge.
- Stansbury, J., & Barry, B. (2007). Ethics programs and the paradox of control. *Business Ethics Quarterly*, 17(2), 239–261.
- Staub, E. (1999). The roots of evil: Social conditions, culture, personality, and basic human needs. *Personality and Social Psychology Review*, 3(3), 179.
- Stellar, J. E., Manzo, V. M., Kraus, M. W. et al. (2011). Class and compassion: Socioeconomic factors predict responses to suffering. *Emotion*, 12(3), 449–459. doi:10.1037/a0026508
- Taylor, C. J. (1982). The diversity of goods. In B. Williams & A. Sen (Eds.), *Utilitarianism and Beyond* (pp. 129–144). Cambridge: Cambridge University Press.
- Thoroughgood, C. N., Sawyer, K. B., Padilla, A. et al. (2016). Destructive leadership: A critique of leader-centric perspectives and toward a more holistic definition. *Journal of Business Ethics*, 151(3), 627–649. doi:10.1007/s10551-016-3257-9
- Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, 64, 231–255. doi:10.1146/annurev-psych-113011-143812
- Tönnies, F. (1887/2001). *Community and Civil Society* (J. Harris & M. Hollis, Trans.). Cambridge: Cambridge University Press.
- Treviño, L. K., Weaver, G. R., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management*, 32(6), 951–990.
- Triandis, H. C., Bontempo, R., Betancourt, H. et al. (1986). The measurement of the etic aspects of individualism and collectivism across cultures. *Australian Journal of Psychology*, 38(3), 257–267.
- Umbrello, S. (2019). Beneficial artificial intelligence coordination by means of a value sensitive design approach. *Big Data and Cognitive Computing*, 3(1), 5. doi:10.3390/bdcc3010005
- Uys, T. (2022). *Whistleblowing and the Sociological Imagination*. New York: Palgrave Macmillan.
- Vianello, M., Galliani, E. M., & Haidt, J. (2010). Elevation at work: The effects of leaders’ moral excellence. *The Journal of Positive Psychology*, 5(5), 390–4114.
- Voelkel, J. G., & Feinberg, M. (2017). Morally reframed arguments can affect support for political candidates. *Social Psychological and Personality Science*, 9(8) 917–924. doi:10.1177/1948550617729408
- Walker, L. J., & Frimer, J. A. (2007). Moral personality of brave and caring exemplars. *Journal of Personality and Social Psychology*, 93(5), 845.
- Walker, L. J., & Frimer, J. A. (2009). Moral personality exemplified. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 232–255). New York: Cambridge University Press.
- Walker, L. J., Frimer, J. A., & Dunlop, W. L. (2010). Varieties of moral personality: Beyond the banality of heroism. *Journal of Personality*, 78(3), 907–942. doi:10.1111/j.1467-6494.2010.00637.x
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, 86(4), 629–647. doi:10.1037/0022-3514.86.4.629

- Walker, L. J., & Pitts, R. C. (1998). Naturalistic conceptions of moral maturity. *Developmental Psychology*, 34(3), 403–419.
- Walsh, C. (2001, October 1). Reconstructing Larry: Assessing the legacy of Lawrence Kohlberg. *HGSE News*. Retrieved from http://www.gse.harvard.edu/news/features/larry10012000_page1.html
- Weber, J., & Wasieleski, D. (2001). Investigating influences on managers' moral reasoning. *Business & Society*, 40(1), 79. Retrieved from <http://bridge.liblink.umn.edu/bridge?genre=%2F%2F%2F%2Farticle&sid=ProQ%3Amaterialsbusfile&atitle=Investigating%20Influences%20on%20Managers%27%20Moral%20Reasoning&title=Business%20%26%20Society&issn=00076503&date=2001-03-01&volume=40&issue=1&spage=79&au=Weber%2C%20James%3BWasieleski%2C%20David&pqdiss=&rft.format=journal>
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136.
- Zapko-Willmes, A., Schwartz, S. H., Richter, J. et al. (2021). Basic value orientations and moral foundations: Convergent or discriminate constructs? *Journal of Research in Personality*, 92. doi:<https://doi.org/10.1016/j.jrp.2021.104099>
- Zimbardo, P. G. (2004). A situationist perspective on the psychology of evil: Understanding how good people are transformed into perpetrators. In A. G. Miller (Ed.), *The Social Psychology of Good and Evil* (pp. 21–50). New York: Guilford Press.

Part II

Influences

Final Proofs

Final Proofs

Personality

Socrates: And temperance makes them temperate?

Protagoras: Certainly.

(Plato, *Protagoras* [332b])

Socrates and Protagoras are arguing here about the nature of virtuous character and whether it can be taught. One intermediary conclusion is that it is virtue that makes one virtuous. Popular advice literature often appeals to the *character* of the individual that primarily motivates and constrains his or her moral action (e.g. she has integrity, he is introverted). Character is viewed as a thing or a property one possesses. In this “character as trait” approach, the moral, virtuous, person acts the way they do primarily *because of the character* they have.¹ We will quickly find out that moral action from the perspective of personality theory is more complicated than this.

We present a picture of moral personality that is *adaptive* to its surroundings, *pluralistic* in its possible forms, and *integrative* of personal and situational influences. Under the integrative heading, we look at three different models that each explore personality as the answer to the question, “What does a person like me do in a situation like this?” When we have finished constructing this more complex picture of what one might mean by moral personality, we then review the empirical literature on personality and moral action using a three-level organization proposed by McAdams and Pals (2006): dispositional traits (e.g. conscientiousness), characteristic adaptations (e.g. authoritarianism), and self-defining life narratives (unique for each individual but with shared structure). We conclude with the claim that it is not our virtues that make us virtuous but rather it is the complex workings of personality that lead us to behave in ways we call virtuous, and that there is no single form this virtue takes.

¹There are even places in academic literature that sound at first brush as though they assume this simplistic model (Doris, 2002; Hartshorne & May, 1928; Hartshorne, May, & Maller, 1929; Hartshorne, May, & Shuttleworth, 1930; Helzer et al., 2014). Some of the “search for character” literature rarely takes time to ask *how* character is supposed to relate to behavior.

4.1 Beyond Character as Traits

Our approach to moral personality is *adaptive* in that it recognizes that personality adapts to the context in which the person lives. It is *pluralistic* in that, unlike earlier approaches (Colby & Kohlberg, 1987), there is no single definitive “moral character” or even set of virtues. It is *integrative* in that aspects of the other influences (moral identity and skills and knowledge) need to be incorporated into our understanding of personality. In this way, personality is not a separate influence on moral action but in some way an integration of the other two influences.

4.1.1 Personality as Adaptive

The “character as trait” model that early psychologists used might lead one to expect what statisticians would call strong main effects for personality characteristics: that character should primarily determine behavior and that this effect should not be affected by changes in the situation (or in any other variable). However, most researchers now agree that one should not expect such straightforward effects: the effect of personality dispositions should be complicated by, at the least, situational and role constraints and social norms (Penner & Orom, 2010, p. 17).

This means we should look for those things that structure the interaction between (1) the character of the individual and (2) the situation in which character is (or is not) expressed. Penner and Orom (2010) suggest that to the extent that the moral decision process is more thoughtful and planned, personal dispositions (though not necessarily reason) are more likely to influence decisions to engage in pro-social behavior. Work in self-regulation (e.g. Hofmann, Friese, & Strack, 2009) also shows that impulsive action is more under the control of the situation than planned action but that even impulsive action can be influenced by the individual taking precautions to remind themselves of their commitments (Gollwitzer et al., 2011).

Still, even very strong situations do not always swamp dispositional characteristics. Coverage of Milgram’s famous (1963) obedience experiments often focuses on the two-thirds of subjects who showed obedience.² But Milgram was just as interested (if not more) in the one-third that managed to avoid the power of the situation, and he included detailed interviews with these individuals in his (1974) book. Penner and Orom (2010) provide another example of a strong situation that fails to influence many who might be disposed to nurturing behavior: parents’ sympathy for their sick children. They provide evidence that non-empathic parents show less empathy and social support for their children when the children are undergoing cancer care. This surely counts as a strong situation that should overwhelm personal dispositions, but the dispositional tendencies show an effect despite the situation.

²And if you read Milgram’s (1963, 1974) descriptions more carefully, even the obedience was not mere passive submission. For most, it was suffused with protest, negotiation, reluctance, covert sabotage, and regret. For some it was stoic persistence in the face of a difficult task. For (some few) others, it was ready agreement. See Browning’s (1992) historical study of a German police battalion during the Holocaust for a similar range of responses to a very strong situation.

Matsuba, Hart, and Atkins (2007) present a mediational model that suggests that the effects of personality³ on volunteering are mediated by two factors: the *moral identity* of the individual (do they think of themselves as helpful?) and the *opportunity* available to them (are they in contact with organizations and people who can connect them to volunteering possibilities?). They found that under- and over-controlled individuals had less salient helping identities, and these individuals were not very affected by the presence of opportunities to volunteer, in comparison to those resilient individuals whose volunteering *was* influenced by opportunity. So here we have a situation (volunteering opportunity) that is weak in influencing behavior, except when there is some match between the personality of the individual and what the situation calls for.

Mediational models like this have led us to think that some way of combining the dispositions of the individual with their experience of the situation should help us better understand the dynamics that lead individuals to see a particular situation as one that calls for them to behave in a particular way.⁴ Models like these have emerged in at least three unconnected literatures. Walter Mischel, the prototypical personality iconoclast, has developed a research program that shows that one can find consistency in behavior if one looks for consistency in the pattern (or “behavioral signature”) of how the individual responds to situations they construe as relevant (Mischel, 2004; Shoda, Mischel, & Wright, 1994). We will review this work in Section 4.1.3.2. Weber, Koppelman, and Messick (2004) have reviewed the literature on how people make choices in social dilemmas (where one individual’s private interests are pitted against shared interests). They too have found that individuals respond as though they had asked “What does a person like me do in a situation like this?” – a question that integrates personality and situation (Weber et al., 2004, p. 281). Gustafson and Mumford (1995) conducted a study with leaders in the US Navy that shows an even more complex personality–situation interaction. An individual’s pattern across several personality variables (called personal style) could be matched with five different empirically derived patterns of situation to predict the leader’s success in those situations. This is another case, in an independent area of research, of how one can find consistency if one takes both the personality and the situation into account. It also suggests that there is no single type of “best leader,” but multiple types. Indeed, current models of leadership see it as a “dynamic cocreational process among leaders, followers, and environments” (Thoroughgood et al., 2016, p. 627). We cover this approach to leadership in a bit more detail in Chapter 3.

We will find a similar personality pluralism in the literature on moral psychology, leading to the conclusion that there is no single “best character.” It is apparent that cultures, organizations, and even friendships are not homogenous. Diversity of goals, values, personality characteristics, skills, etc. has long been recognized in research, but

³In their case measured as resilient vs. under- and over-controlled personality. Resilient personality was measured in children aged eight to ten years before volunteerism was measured (Atkins, Hart, & Donnelly, 2005). In children, a resilient personality means they are compliant, socially skillful, and have positive emotional tone. Under-controlled children have low levels of impulse control and trouble maintaining relationships. Over-controlled children are shy, timid, anxious, and have trouble maintaining relationships.

⁴See also Chapter 5 for another way to think of this match-to-situation.

rarely investigated on its own terms. It is at least clear that the reaction to moral diversity is often negative, and we judge harshly those who have moral disagreements with us. We need to know more about how people think about the diversity of moral judgments and goals in a moral ecology. What levels and kinds of moral diversity are people willing to tolerate? How do people handle moral tradeoffs when diverse values are in play? How do people choose the moral goals they strive to achieve? How do individuals lobby to turn optional moral goals into mandatory goals? There are surely even more questions than these.

4.1.2 Personality as Pluralistic

In his 1991 exploration of the varieties of moral personality, the philosopher Owen Flanagan uses a short catalog of Christian saints as an example of one sort of morally exemplary person. He then shows that even this somewhat limited category of moral excellence contains multitudes. Saints are introverted and extraverted, patient, irascible, thoughtful, simple, impulsive, warriors, poets, scholars, poor, rich, etc. The saints Flanagan refers to have in common the requirements of the Roman Catholic Church but very little else.⁵ From this example, one could expect that psychological research on moral excellence will also find multiple models.

In their review of the implications of personality psychology for moral psychology, Hill and Roberts (2010, p. 380) include as their first proposition that “A singular moral personality does not exist.” One would expect this from the example of Roman Catholic saints, and research bears it out. Work on moral personality often proceeds along the lines of identifying individuals who are remarkable for their moral commitment and then attempting to learn something from their histories and personality profiles.

This tradition begins with the work of Oliner and Oliner (1988) who studied individuals who participated in rescue attempts of victims of the Holocaust. They found three primary motivations for rescuers: those who participated because of their connection to groups who encouraged it (called “normative,” 52% of the rescuers), those who responded to empathy for particular individuals (“empathic,” 37%), and those who participated because of principled decision (“principled,” 11%). Colby and Damon (1992) also sought out moral exemplars but in the field of social service in the United States. Their extensive nomination method resulted in interviews with twenty-three moral exemplars who differed widely in education and the targeted goal of their service (e.g. medical service for the poor, entrepreneurship, civil rights, urban poverty, etc.) They noted that the individuals they interviewed tended to specialize in one of two types of work: social reform or direct helping of those in need. In a study of moral exemplars in the field of computing, Huff and Barnard (2009) documented a similar specialization, with some engaging in societal reform and others using their skill to help others directly. Walker and colleagues (Matsuba & Walker, 2004; Reimer, DeWitt Goudelock, & Walker, 2009; Walker & Frimer, 2007; Walker, Frimer, &

⁵The Roman Catholic Church canonizes saints who are recognized for “heroic virtue” (Beccari, 1907), those “whose virtues and good works greatly outdistance those of ordinary good people.”

Dunlop, 2010) have done the most definitive work in looking for varieties of moral personality among exemplars. They constructed a sample from winners of two humanitarian prizes in Canada, one for heroic action and one for influential long-term volunteer work, and identified at least three types: just, brave, and caring.

If, in the quest for some unifying principle of morality, one moves to catalogs of virtues, one confronts a similar pluralism. Kohlberg and Mayer (1972) have complained that lists of virtues tend to be arbitrary and have only little common overlap. In addition, some of the virtues seem to be at least in tension with one another, e.g. justice and charity. Cross-cultural work on values by Schwartz (2006, 2010) provides some structure to the tension in the domain of values, with a circular model that shows each value in tension with some other value on the other side of the circumplex.⁶

Thus, it seems that in moral personality, motivation, goal, virtue, and value we are faced with a pluralism that is difficult to reconcile with insistence on the unity of the virtues (Walker et al., 2010). The personality of moral exemplars, their motivations, and their goals seem to be widely variable.

4.1.3 Personality as Integrative

We now have an outline of moral personality as adaptive and pluralistic. What about integrative? It is in response to the complexities listed so far that psychologists have begun constructing integrative models of personality. These systems are usually based in a wide review of the empirical personality literature and then constructed with the express purpose of explaining how the individual maintains a sense of coherence and efficacy in their world.

4.1.3.1 Idiographic vs. Nomothetic Approaches All the models we will review in this section are *idiographic*: they look at psychological processes as they occur at an individual level – how each person answers the question “What does a person like me do in a situation like this?” This is in contrast to a standard psychological *nomothetical* approach that looks at aggregates of individuals (e.g. do extraverts on average tend to volunteer more?⁷).

Nomothetic approaches, like those associated with the five-factor model of personality (McCrae & Costa, 1997), look at differences between individuals, averaged over many people, and the relationship of those differences to other things of interest, such as volunteering. These patterns say something about the behavior of abstract groups that are constituted statistically, but it is not immediately obvious how the patterns translate into the functioning of the individual.

The distinction between nomothetic differences measured at the group level and individual, idiographic personality architecture is a difficult one; but one with important implications for the conversation about what character (and virtue) is and how it

⁶See Chapter 3 for a detailed explanation of this work. We argue in Chapter 9 that this tension is part of the basic human condition, and that learning to operate within it is part of what becoming moral is about.

⁷Yes. See Ozer and Benet-Martínez (2006) for evidence of this relationship.

functions. It is certainly too simplistic to say, with Socrates and Protagoras, that “temperance makes men temperate.” And it is too simplistic to say that extraversion or conscientiousness – as defined by current personality inventories (e.g. Costa & McCrae 1997) – is a property one can have. Nor can one easily say that this property causes behavior in any straightforward way. This critique is fully elaborated in Cervone (2004), with the significant implication that one cannot, without empirical evidence, talk about extraversion or conscientiousness as psychological constructs that an individual “has” in the same way one “has” a superego in Freudian approaches (if anyone indeed has a superego).

Conscientiousness (as measured by most personality inventories) could, theoretically, be part of an internal psychological system that influences behavior, and some theorists speak about it this way (McCrae & Costa, 1997). But the evidence in fact, seems to lean against this (Borsboom, Mellenbergh, & van Heerden, 2003; Cervone & Tripathi, 2009). Even though the Big Five personality characteristics (extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience) are useful in explaining differences between individuals in the aggregate, attempts to show they have the same psychological function at the individual level have been, at best, equivocal. Two individuals, for instance, may be much more honest than the general population and both may tend to tell the truth across many situations. Still, those situations will be different for each person and each individual may be motivated to be honest for different reasons in the different contexts (sometimes having nothing to do with a desire for truth).

This distinction is relevant to the current debate in philosophical virtue theory about the existence of character (Doris, 2002; Snow, 2010). A conceptual virtue such as honesty or generosity may well describe the differences among groups of people but describe neither the reasons or patterns of honesty in any individual person nor the complexity and interdependence of that virtue with another virtue, for that person. And thus, our naïve psychology that honesty is a thing one “has” may mislead us to expect a kind of consistency where, for good reason, there should be none. The nomothetic approach may be correctly answering the question of how people, in the aggregate, are different from each other, but not giving us an answer about how individuals, as they go about their lives in the world, constitute those differences.

Both Mischel (2004) and Cervone and Tripathi (2009) ask this latter question. How does the individual as a whole act in the world? Both approach individual functioning by calling on psychological systems of cognition and emotion that underlie and help to produce the individual’s characteristic responses. Thus, according to Cervone (2004, p. 430), “... an intraindividual focus does not indicate disinterest in individual differences. Instead, it represents a strategy for understanding them: a bottom-up strategy ...” that articulates how each person is honest in the characteristic way he or she is.⁸

⁸A cynical, but partially true, implication of this is that the entire personality-situation controversy that has roiled psychology and philosophy can be reduced to a measurement issue (top down vs. bottom up). There have been amazing advances in measurement at the idiographic level (Beck & Jackson, 2020). But more reasonably considered, it has been a long standing theoretical *and* empirical puzzle that we are only now beginning to unravel.

Because they focus on the psychological processes that support the adaptiveness and coherence of individual action, the models we review here are particularly relevant to thinking about how people act morally. It is interesting, and perhaps useful for recruiting purposes, to know that an identifiable type of person (e.g. extraverts) volunteers more often. This is the nomothetic level. But if we want to understand how and why each one volunteers, the structure and functioning of individual character, we need to look to psychological models of the individual, idiographic “springs of action.”⁹

4.1.3.2 The Cognitive-Affective Processing System Mischel and colleagues’ (Mischel, 2004; Shoda et al., 1994) work on *personality signatures*, or individual patterns of responding, is primarily designed to try to understand how individuals navigate different situations with a sense of coherence. They offer a powerful and empirically well-supported answer, and all the other integrative models mostly propose parallel answers in response to this central puzzle of personality. They begin with noting the two types of personality theory that we have just covered.¹⁰ They then set for themselves the task of determining how to measure idiographic consistency based on “*intraindividual patterns of variability*” (Mischel, 2004, p. 9, italics added), the distinct pattern each individual shows in response to perceived situations.

Shoda, Mischel, and Wright’s (1994) work, based in detailed tracking of children at a summer camp, provides evidence that one can collect data to construct these behavioral signatures and that these signatures have predictive value. They followed aggression and self-regulation in 84 children across the 6 weeks of the summer camp, with an average of 160 hours of behavioral observation of each child over that time. They collected videotaped social interactions and coded them, coded the children’s activities within each hour for behaviors like whining, compliance, and physical and verbal aggression, and asked camp counselors to give global personality judgments for each child. This required a team of 77 adult observers, with each child observed by between 14 and 28 observers over their time at camp. This produced an enormous and complicated database of each child’s actions in strategically selected situations.

They were able to show striking stability in the behavior profile of each child across five different types of situations (peer tease, peer approach, adult praise, adult warn, adult punish). Thus, individuals who might on average seem similarly aggressive, nevertheless showed clear *differences* across the situations that called forth that aggression. Some children showed clear stability in these differential reactions to each situation, while others were more variable. One child, for instance was only aggressive in response to adult punishment, while another was most aggressive to peers when they initiated positive contact (peer approach).

⁹The origin of this phrase comes from the philosopher Jeremy Bentham’s 32-page pamphlet *Table of the Springs of Action* (1817). It consists of 14 tables showing the origins of motivations, accompanied by copious notes.

¹⁰Mischel (2004) calls nomothetic consistency “Type 1” and idiographic consistency “Type 2.”

They conclude that the individuals showed more or less¹¹ stable patterns of variability in their behavior across different types of situations. Mischel (2004) explains these patterns, or personality signatures, with a model of a cognitive-affective processing system (CAPS). This model suggests that individuals differ in their mental representations of “self, people, and situations, enduring goals, expectations-beliefs, and feeling states, as well as memories of people and past events” (p. 11). These representations are not isolated, but are organized in interaction with each other, and in characteristic patterns of emotion and cognition. Everyone’s unique patterns of behavior are produced by these unique representations, processed in characteristic ways by the individual. They have tested the model with computer simulations (Shoda, LeeTiernan, & Mischel, 2002), and shown that it can produce both nomothetic (on average) and idiographic consistency.

It is important at this point to remember that we are so far not talking about *moral* consistency. The consistency in personality does not need to be one of moral consistency. There are a variety of identity themes (McGregor & Little, 1998), or personal projects, that individuals construct to organize their lives, including being a meaningful part of a community, achieving status or success, or simply having fun. Indeed, in a study by McGregor & Little (1998) that asked over 300 college students about their personal projects, moral ideals did not emerge as a major theme among any significant subset of the respondents. One is not guaranteed that the consistencies in life will be moral ones. Commitment to simple hedonism, achievement, or group loyalty may even result in one being labeled a moral wanton (Frankfurt, 1971) but not, perhaps, a wanton in respect to one’s constructed identity theme (e.g. having fun or being dominant). One need not be labeled immoral because one does not take morality as a central aspect of the self. See the chapter on Self for the complexities of how we construct moral identities. We expect minimal commitments to moral principles in order to construct a habitable society (Haidt, 2008a; Haidt & Joseph, 2004; Tomasello & Vaish, 2012), and can stay within these lines without having morality at the center of one’s identity.

To conclude, Mischel and colleagues’ work (Mischel, 1968, 2004; Shoda et al., 2002; Shoda et al., 1994) has provided a paradigmatic solution to what has been called the “personality paradox”¹² and helps us understand how a person might be empathic and helpful in one situation but not in another. But there is more to the role of personality in moral action than the personality paradox. Individuals strive to become better over time (or not). They try, for instance, to become more patient in their relationships or more courageous in their work life. Individuals’ behavioral signatures show consistency (or lack thereof) depending in part on their self-regulatory skill. They resist (or not) the temptation to make a snarky comment and they resist (or not) the desire to impress others by embellishing a personal story. They construct personal narratives that support a career of helping or personal narratives that undermine it. In response to events in the

¹¹ Some would show stability for one behavior (e.g. aggression) across situations, but not for another (e.g. whining). The complications here are immense, but one can find pattern in them.

¹² This is the lack of relationship between what a “personality characteristics as traits” approach would expect in behavior consistency, and what we actually find.

world, they construct worldviews that emphasize cynicism or charity. In addition to making CAPS-like movements to resolve the personality paradox, the personality systems we review next also try to address the broad spectrum of personality's influence on moral action.

4.1.3.3 Personality Architecture (KAPA) Daniel Cervone and his colleagues (Cervone, 2004; Cervone & Tripathi, 2009) provide an alternative model with similar characteristics to CAPS. They refer to their theoretical system as “architecture” to distinguish it from the nomothetic approach so prevalent in recent personality psychology. Cervone's (2004) Knowledge and Appraisal Personality Architecture (KAPA) model is similar to Mischel's in that both refer to the ways individuals construe their world and react to it. Cervone's model helpfully connects this to a large literature in emotion¹³ that helps us understand emotions as mixtures of cognition (appraisals) and reactions to those cognitions. Because individuals have different conceptions of who they are, of who their friends and enemies are, and of what harm or help mean, they will have different reactions to what might objectively look like the same action. The research on emotional appraisal (Lazarus, 1991) helps us to understand, from the bottom-up, these different reactions, and the KAPA system (Cervone, 2004, 2021; Cervone & Bartoszek, 2013) helps us to see the characteristic ways of reacting as a part of our personalities, and how they might change over time.

4.1.3.4 The Neo-Socio-Analytic Model (NSAM) Roberts and Wood (2006) propose an integrative model that, like those of Shoda and Mischel (1994) and Cervone (2004), uses an idiographic approach to resolve the personality paradox, but reaches much further in its integration and in its goals. It is a complicated model that considers influences ranging from genetic mechanisms to society and culture.¹⁴ For our purposes, we will emphasize several similarities between this comprehensive model and the approach taken in this book. Like the framework in this book, the NSAM sees the social surround of the individual as crucial in shaping the individual's pattern of response. We cover this aspect in Chapter 3. Roberts and Wood's (2006) model also, like this book, emphasizes the role of abilities, skill, and knowledge in shaping the characteristic ways the individual as a whole acts in the world based on the intersecting influences of the many parts of the model.¹⁵ The resulting model is more comprehensive than others, but also more cumbersome. Hill and Roberts (2010, p. 382) use this model to make specific claims about moral psychology and personality, among them the claim

¹³For a detailed discussion of emotion, see Chapter 8, along with discussions of emotional regulation in Chapters 6 and 9.

¹⁴The complete list of influences consists of the following: genes, physiological mechanisms, Big Seven personality traits, positive and negative affect, attachment styles, motives, values, goals, interests, life tasks, abilities, general intelligence, verbal and spatial ability, quantitative skill, stories, significant memories, scripts, ideological settings, identity, reputation, social roles, and social and cultural influence.

¹⁵See also Matthews (2008) for another approach that emphasizes skills.

common to all the models we have reviewed here: “Studies of moral personality development must go beyond just traits.” Recently one aspect of the model has been specified in more detail: how short term, situational processes can produce longer term personality change (Wrzus & Roberts, 2017). We review this advancement in Chapter 6.

4.1.3.5 A Three-level Approach to Personality Systems The final approach to personality we present here is more of a framework than a theoretical model of individual personality processes. But the framework organizes a broad swath of literatures and suggests how they might relate to each other. McAdams (2009, p. 12) proposes that personality consists of “(1) an individual’s unique variation on the general evolutionary design for human nature, expressed as a developing pattern of (2) dispositional traits, (3) characteristic adaptations, and (4) self-defining life narratives, complexly and differentially situated in (5) culture and social context.” Note that this approach centers, again, on the individual, and like Roberts & Wood (2006) encompasses biological roots and cultural surround. The three main parts are conceived of as levels of increasing specificity for the individual, with each level providing a more nuanced understanding of the individual. *Personality traits* are things like extraversion, dominance, and emotional reactivity, and they account for some broad consistency of individuals across time and situation.¹⁶ *Characteristic adaptations* fill in the details of the individual’s reactions in situations. This level includes the schemas and emotional appraisals one finds in Mischel’s (2004) and Cervone’s (2004) systems, along with things like motives, goals, strivings, values, coping strategies etc. This level of description helps provide articulation to the broad patterns outlined at the level of personality traits. In the same way that people vary on traits like extraversion, one can arrange people into groups based on their values (Graham, Haidt, & Nosek, 2009; Graham et al., 2011; Haidt, Graham, & Joseph, 2009; Schwartz, 2006, 2010) or motivations (Snyder & Clary, 2000). The most specific level of personality is that of the individual’s *narrative* of their past and expected future. These narratives are, of course, unique, though one can find some shared structure in them. For instance, when they tell stories about their life, some people tell stories whose theme is often about how good outcomes flow from bad occurrences, e.g. how losing a job led to a new step in life. The tendency to tell stories with these “redemption” themes has implications for one’s moral profile.

4.2 Personality Influences on Moral Action

The remainder of this chapter uses this three-level framework by McAdams and Pals (2006) to organize a review of how each level is related to moral action. As you read, remember to keep the nomothetic *vs.* idiographic distinction in mind: is the pattern

¹⁶They “account for” this consistency in that they point out patterns of consistency, but they may not be the actual “springs of action” that produce the consistency in any individual (Cervone, 2004).

the research describes just based on abstract categories, or are the categories things that one might expect to find having real influence at the individual level?

4.2.1 Dispositional Moral Traits

The early work of Hartshorne and May (1928) and the conclusions of Mischel (1968) have bolstered the claims in philosophical circles (Doris, 2002; Doris & Stich, 2005) that large-scale personality traits such as the Big Five do not help us understand moral action. But recent work has, in fact, found reliable effects of these characteristics in areas such as helping, cooperation, criminal behavior, and espoused moral values. In this section we review this work and conclude with Dovidio et al. (1991, p. 101) that “despite the pessimism of earlier reviews in this area, a growing body of literature suggests the importance of individual differences in helping.” These effect sizes are generally small, but this is what one would expect in a science where there is massive interaction among the causal variables (Götz, Gosling, & Rentfrow, 2022). The section begins with a review of the correlates of each of the Big Five personality traits, then notes the involvement of moral themes in alternative six- or seven-dimensional formulations, and finally notes some cross-cultural attempts to present virtues as personality-style traits.

4.2.1.1 Moral Correlates of the Big Five The Big Five is the standard name for the five-dimensional outcome of a particular approach to personality traits. These psychological constructs do a remarkably good job of covering the broadest span of applicable personality traits (John & Srivastava, 1999; Roccas et al., 2002). There is, of course, continuing conversation about their comprehensiveness, their causal and ontological status, and their structure (Cervone, 2004; John & Srivastava, 1999). The dimensions have been shown to replicate reasonably well in Western cultures and moderately well (with informative exceptions) in non-Western cultures (McCrae & Costa, 1999; McCrae et al., 1998). They are usually conceived as consisting of five dimensions (e.g. introversion–extraversion), with each dimension being a broad level summary of more specific traits (e.g. for extraversion, sociability and dominance). Thus, as one tries to parse the relationships of each dimension to morality, it is good to keep in mind that it may be one aspect of a dimension (e.g. dominance) that is pulling the weight in a correlation (with e.g. brave moral action).

Extraversion–introversion: People differ in the extent to which, across a broad range of situations, they tend to be socially outgoing or dominant. These differences are part of a pattern of interaction with the social world that includes interaction that we would call moral. In their study of moral exemplars who won Canadian prizes for bravery or social service, Walker et al. (2010) found that the brave exemplars were more likely to score highly on social dominance (an aspect of extraversion), while the caring exemplars tended to score more highly on nurturance (an aspect of agreeableness).¹⁷ Huff and Barnard (2009) found that, among the moral

¹⁷See Chapter 9 for a detailed presentation of this research.

exemplars in computing they studied, the more extraverted exemplars were inclined to be involved in social change movements, while some of the more introverted exemplars tended to use their craft to help individuals. Extraversion is also associated with community involvement and volunteerism (Ozer & Benet-Martínez, 2006), although Matsuba et al. (2007) have shown that this relationship can be a complicated combination of motivation and opportunity. Aspects of introversion are also related to how easily one learns the moral structures of one's childhood. Kochanska and colleagues (Kochanska & Aksan, 2006; Kochanska & Murray, 2000; Kochanska et al., 2010) have shown that shy, fearful children are more readily socialized, particularly with regard to the development of feelings of guilt. Thus, the dimension of introversion–extraversion has a complicated relationship to morality, and it might be best to say the dimension may more describe the different ways that individuals are moral: individuals on different ends of this dimension can be “differently moral” rather than more moral.

Agreeableness–disagreeableness: This dimension of personality has variously been called social adaptability, likability, friendly compliance, and love (John & Srivastava, 1999). These different names for the dimension give some idea that its core is about positive social interaction, including aspects of cooperation and empathy for others. Those who score high on this dimension are more cooperative with others, volunteer more (Dovidio et al., 2006) and donate to charities more (John & Srivastava, 1999), while low agreeableness is associated with criminal behavior (Ozer & Benet-Martínez, 2006). Individuals high on agreeableness tend to endorse the values of benevolence and tradition, and to reject power as a value (Roccas et al., 2002). One might conclude that this is a dimension with a moral loading, that is, that as individuals increase in their score on agreeableness, they are also likely to increase in morality. But this would be too hasty. At the extremes of criminal behavior this may work, but one can certainly call to mind famously grumpy and uncooperative saints (Flanagan, 1991) and social action often involves being uncooperative to the point of lawbreaking (Colby & Damon, 1992). Agreeableness may well itself contain two different kinds of morality since Hirsh et al. (2010) have found that the compassion aspect of agreeableness correlates with politically liberal values, while the politeness aspect is associated with conservative values.

Conscientiousness–unconscientiousness: High scores on this dimension are reliably related to success in work environment and career, to impulse control and delay of gratification, and to school and college grades (John & Srivastava, 1999). The orderliness aspect of conscientiousness correlates with what Haidt and colleagues (Haidt & Joseph, 2004) have called the “binding factors” of morality: dedication to in-group loyalty, recognition of authority, and honor of sacred things (Hirsh et al., 2010). Low conscientiousness is associated with criminal behavior (Ozer & Benet-Martínez, 2006) and high scorers are more likely to endorse the values of achievement and conformity (Roccas et al., 2002).

Neuroticism–emotional stability: Neuroticism is the awkward label for the dimension describing an individual's negative emotional reactivity. It is associated with a host of negative outcomes in life, including negative family and work satisfaction, lack of financial security, failure in the workplace, and conflict in romantic relationships. High scores on this dimension are also associated with criminal behavior. Those

who are high in extraversion and low on neuroticism are less emotionally responsive to negative feedback, and more likely to look on the bright side of things. Not surprisingly, low neuroticism is related to better emotional regulation. All these relationships are reviewed by Ozer and Benet-Martínez (2006).

Openness–closedness: Openness to experience is the tendency to seek new experience, to be tolerant of and welcoming to other views (Ozer & Benet-Martínez, 2006). In his review of the moral personality, McAdams (2009) notes that high openness tends to correlate with post-conventional moral reasoning, and low openness with authoritarianism; and on this basis suggests that openness may “be most closely associated with moral reasoning and thought” (p. 15). The empirical relationships are clearly there, but one should be wary of concluding too much about the dispositional nature of morality from them. Haidt and colleagues’ (Graham et al., 2009; Haidt, 2007, 2008a, 2008b; Haidt & Graham, 2007; Haidt et al., 2009; Haidt & Hersh, 2001) work on the dimensional nature of moral cognition suggests that openness is only tracking those two (justice and care) moral dimensions that are accepted by political liberals, and that it leaves conservatives’ moral commitments (more related to group loyalty, obedience, and sacredness) out of the domain of the moral. Some evidence for this assertion comes from the negative correlation openness shows with Schwartz values of conformity, security, and tradition, items one would expect to be related to conservative values (Roccas et al., 2002). Thus, the dimension tracks moral disagreement on which values are more important.

Alternatives to the Big Five: There are several alternatives to the Big Five that use the same dimensional approach but find that six or seven dimensions better fit the way traits cluster together. The approaches we mention here differ from the Big Five in how they handle issues of morality. Both add explicitly moral dimensions to the standard dimensional analysis. One approach (Ashton & Lee, 2007) shows six dimensions that slightly reshuffle aspects of some dimensions (e.g. moving angry emotionality from the neuroticism dimension to the agreeableness dimension). The largest difference is the appearance of a sixth dimension called Honesty/Humility that captures differences based on sincerity, fairness, honesty, humility, and lack of greed. The original Big Five themselves have some moral tone (e.g. conscientiousness, agreeableness) and it is interesting that the additional dimension that arises in this approach is an explicitly moral one. Saucier (2003) reports on a seven-dimensional approach, and again, one of the added dimensions is explicitly moral: Social (Un)acceptability (e.g. rude, greedy, corrupt *vs.* loyal, dependable, honest). These additional explicitly moral dimensions lead one to think one might map the virtues themselves as traits, and there are researchers who have done so.

4.2.1.2 Mapping the Virtues as Traits One of the difficulties with talking about the virtues is that generated lists of virtues are inevitably idiosyncratic, often non-overlapping, and difficult to harmonize (see Lapsley & Narvaez, 2006 for a review of this difficulty). One approach to this problem is to empirically establish groupings of virtues from some suitably large set. Both Cawley (Cawley III, 1997; Cawley III, Martin, & Johnson, 2000) and De Raad (De Raad & Van Oudenhoven, 2011) have

generated extensive lists of virtue relevant trait words, and asked people to rate themselves and others on these traits.¹⁸ The approach, then, is very much like those that have generated the Big Five personality dimensions and its variants. Thus, they tend to find *dimensions on which the virtues themselves vary* (e.g. virtues rank high or low in the extent to which they involve empathy), rather than give us a comprehensive set of virtues. The methodology asks the question: “In what pattern do these traits hang together in people’s self-reports?” Cawley et al. (2000) found a four-dimensional pattern in American English speakers (empathy, order, resourcefulness, and serenity), while De Raad and Van Oudenhoven (2011) found six dimensions in Dutch speakers (sociability, achievement, respectfulness, vigor, altruism, and prudence). Both patterns showed clear relationships to the Big Five personality characteristics and some convergence on shared dimensions (De Raad & Van Oudenhoven, 2011). For instance, individuals high on the virtue dimension “order” in Cawley’s (1997) system, tend to score high on De Raad and Van Oudenhoven’s (2011) achievement virtue dimension, and also generally score high on the Big five characteristic of conscientiousness.

A more classic approach is that of Peterson and Seligman (2004) who produce a catalog of important virtues and their supporting strengths. They survey a wide range of cultural, philosophical, and psychological approaches and argue that their list of six virtues (wisdom, courage, humanity, justice, temperance, and transcendence) is at least implicit in the approaches they survey. They argue, for example, that though humanity is never explicitly mentioned in Plato’s or Aristotle’s account of the virtues, “notions of shared humanity ... are scattered across both works” (Peterson & Seligman, 2004, p. 47).¹⁹ Most of the effort is expended on the construction of a set of “strengths” that support each virtue. The virtue of humanity, for example, is composed of the strengths of love, kindness, and social intelligence. This conceptual framework is the basis for an extensive psychometric effort to produce self-report and observational scales for each of the twenty-five strengths they catalog.

Lapsley and Lasky (2001) take a cognitive approach to untangling the puzzle of the structure of the virtues. They propose that our understanding of the virtues is a “prototype,” a cognitive network with links based on similarity among the items. Walker and Hennig (2004) base their research into folk conceptions of morality on this cognitive network approach, and in places Peterson and Seligman seem to do so as well (Peterson & Seligman, 2004, p. 16). As one might by now expect, Lapsley and Lasky (2001) generated a large list of character-relevant traits, and had people rate how much each applied to someone “of good character.”²⁰ They then

¹⁸For instance, De Raad and Van Oudenhoven (2011) started with 1,203 trait descriptive words in Dutch.

¹⁹They do not mention that both Aristotle and Plato had quite limited views of how widely full humanity was shared: women, slaves, foreigners (barbarians), etc. were excluded (see Nussbaum, 1996, for a valuable and nuanced discussion of virtue in Hellenistic philosophy).

²⁰They began with 175 traits generated by subjects. At the end, the top five character-relevant traits were honest, trustworthy, genuine, loving, and dependable. The bottom five, least character-relevant traits were well-noticed, provocative, shy, popular, and lucky.

showed that one could produce the expected cognitive effects that indicated the traits were organized as a prototype. For instance, in a task where they first read trait words associated with a person, and were then asked to recognize those words, they falsely remembered character-relevant words that had not been shown. This and other effects suggest these character-based traits are part of a cognitive network based on similarity, with some traits being more central to the network and others more peripheral.

4.2.1.3 Some Conclusions About Traits The most obvious outcome from this overview of large-scale personality traits is in harmony with the conclusion by Hill and Roberts (2010) that there is no singular “moral personality.” There are many different personality traits that are related to morality and moral action, and there seems to be no clear unity to them. In Chapter 7 we will review evidence (Schwartz, 2010) that the values that underlie many of the moral aspects of personality may even have a structure that pits them against each other, with any unity lying more in a sense of appropriate balance for the situation than with clear compatibility of the values or personality traits that support them.

Thus, some version of value pluralism reigns in the domain of moral personality. There are many types of moral exemplars and multiple personality profiles that can each legitimately be called moral. No particular dimension of the Big Five captures the variety of ways people are moral. When one views virtues as though they are traits, these traits do not naturally cohere, and no single virtue trait captures the variety of ways people are virtuous.

But this state of affairs is exactly what one would expect if moral personality is *adaptive* to its surroundings, *pluralistic* in its possible forms, and *integrative* of a variety of influences. Effects at the individual trait level should be small and interactive with many other variables (Götz et al., 2022).

Finally, the psychological status of traits or virtues is unclear. Certainly, people use trait/virtue words to describe themselves and others, and a large literature exists that shows some structure in the ways these descriptions cohere within a limited set of dimensions. But it seems unlikely that these dimensions refer to specific psychological entities (in the way that motives, values, goals, memories, scripts, self-concept, emotions, intuitions, etc. do). It is likely that some combination of the entities just listed influence the behavior that we then label with trait/virtue words. These labels are likely based in prototypes, cognitive networks of descriptors whose links are based on similarity. Our experience of others’ behavior, and our culture and language, are the basis of our folk conception of the virtues. This conception is then manifest in our prototypes of patterns of moral action (Lapsley & Lasky, 2001; Walker & Hennig, 2004). Current models of virtue (Fowers et al., 2020) or character (Narvaez, 2010) have incorporated the *adaptive, pluralistic, and integrative* nature of moral action and provide an initial structure to grasp the complexity.

To understand how we become moral, learn the virtues, and take moral action, we will need to look more carefully at the “springs of action” that produce these patterns. Some of these springs flow from McAdams and Pals’s (2006) second level of personality, characteristic adaptations.

4.2.2 Characteristic Moral Adaptations

This level consists of the characteristic ways that an individual adapts to his or her environment(s). To say “his or her” already suggests that there is likely more than one way of adaptation, and more than one kind of environment. As in the more basic level traits, multiplicity and pluralism will be the rule here. These aspects of personality differ from the broad trait aspects mentioned earlier in that they are more:

- closely linked to motivation and cognition
- easily influenced by environment and culture
- likely to change over time (through therapy, practice, or environmental influence)

Roberts and Wood (2006) suggest that these lower-level aspects have more clear implications for situational action and might therefore be the mechanisms that produce the patterns of characteristic action at the higher levels (see also Little, 1999; McAdams & Pals, 2006). There is no canonical list of these aspects, but they are at least *motivational* (e.g. life goals and projects), *cognitive* (e.g. self-perceptions), and *emotional* (e.g. dispositional empathy). All likely involve some combination of these. For instance, life goals include some cognitive conception of the good, self-perceptions include motivational aspects, and empathy requires knowledge and motivation.

Many of the items that might be listed in this section of the text are covered in other chapters, in part because the literatures group more easily that way. McAdams (2009), for instance, lists the content of moral beliefs and values and the structure of moral reasoning as two aspects at this level. We review these aspects in detail in Chapters 7 and 8 since these aspects play such an important part in each process. What follows is an incomplete listing of some of these lower-level aspects of personality.

4.2.2.1 Pro-Social Personality There is a long and productive research program in what has been called “pro-social personality” (for reviews, see Dovidio et al., 2006; Penner et al., 2005; Penner & Orom, 2010; Schroeder & Graziano, 2015). This “cluster of interrelated dispositions” (Penner & Orom, 2010, p. 60) is a reliable predictor of ongoing involvement in pro-social behavior, including long-term volunteering, spontaneous helping during accidents and other incidents (in lab studies and in real life), help to friends, organizational citizenship, and parent’s empathic response to their sick children (Matsuba et al., 2007; Penner & Orom, 2010). The multidimensional nature of the cluster suggests, again, that there is more than one single aspect to moral action. The most recent versions of this approach (Penner & Orom, 2010) present a two-dimensional picture, with one dimension of *other-oriented empathy* and a second of *helpfulness*. *Other-oriented empathy* consists of pro-social thoughts and feelings, including thoughts and feelings of social responsibility, empathic concern, affective and cognitive empathy, and other-oriented moral reasoning. This first dimension correlates moderately with the Big Five dimension of agreeableness and with a nurturant or communal interpersonal style (Penner & Orom, 2010). The *helpfulness* dimension consists of a reported history of helping, and a tendency to not experience personal, egocentric distress in response to the distress of others. This later dimension thus includes a positive inclination toward helping and the removal or absence of what might be a barrier to helping others,

one's own egocentric distress. This second dimension correlates positively with agency or dominance, and with feelings of self-efficacy and personal control (Penner & Orom, 2010).

The two dimensions correlate with each other, but not strongly, and it is the combination of the two dimensions, empathy and an action orientation, that predicts helpfulness. Thus, the two dimensions combine to describe those people who experience empathy for others and engage in moral reasoning about their obligations, but who are also socially dominant enough to have opportunities for and the expectation of success in helping (Penner & Orom, 2010 p. 63). Phrasing the two dimensions in this way makes the interaction of personal characteristics with situational aspects obvious.

4.2.2.2 Cynicism It is not surprising that someone who has a negative attitude toward others and suspects them of dealing selfishly and dishonestly will be unlikely to engage in helping those individuals. Cynicism has been most extensively studied in the organizational literature (Dean, Brandes, & Dharwadkar, 1998) but has recently made an appearance in work on ethics in science (De Vries, Anderson, & Martinson, 2006; Mumford et al., 2007). In both literatures cynicism is a reaction to perceived inequity and unfairness, and the source of a disposition not to help others. In science, Mumford (2007) has tracked the exposure of graduate students to ethical violations, and shown that this exposure produces cynicism, which in turn is expressed as a willingness to skirt the rules in self-defense. De Vries et al. (2006) have shown a similar pattern in interviews with funded scientists. Similarly in organizations, cynicism is often a response to poor organizational performance, excessive executive compensation, and harsh layoff policies (Dean et al., 1998). It seems possible that cynicism can be domain specific, limited, for instance, to one's workplace. There is some evidence that it is unrelated to any of the Big Five traits (Haas, 2002).

4.2.2.3 Optimism and Positivity These have been cited as characteristic of moral exemplars (Oliner & Oliner, 1988). Colby and Damon (1992) note this among the exemplars they interview, but also point out some significant exceptions of angry, disappointed individuals struggling for social change. As the Colby and Damon exceptions might lead us to expect, the relationship of optimism and positivity to moral action has a spotty record of replication (Walker & Frimer, 2007).

4.2.2.4 Internal Locus of Control and Self-Efficacy These are both characteristics that have been mentioned as a part of the action-oriented helpfulness dimension in pro-social personality research. Locus of control is the general feeling that one has control in one's life. Self-efficacy is a more domain-specific sense that one can be effective in action in a particular area. Both are tendencies that might urge one to act or keep one from acting. Both are related to the willingness to take action in relevant situations (Midlarsky et al., 2005; Penner & Orom, 2010; Treviño, Weaver, & Reynolds, 2006).

4.2.2.5 Generativity Generativity was defined by Erikson (1963) as having a concern for promoting the well-being of future generations. But the concept has been widened to a more general concern for society (de St. Aubin, McAdams, & Kim, 2004). There

is a developmental arc of adaptation that shows increases in goals related to generativity in mid-life. (McAdams, 2009). One does not have to think of generativity as simply future-generation oriented (Snyder & Clary, 2004), though volunteerism is often present oriented, it does not rule out, and often seeks, future benefit. Generativity is usually construed in positive terms, but it can also be about agency, narcissism, and self-replication (Snyder & Clary, 2004, p. 232). Still, the positive aspect of this goal orientation has been associated with a wide range of pro-social action, including involvement in children's education, responsible parenting, participation in politics and in religion, and in volunteer work for the poor (for reviews, see de St. Aubin et al., 2004).

4.2.3 Narrative Identity

Narrative consists in the way the individual arranges his or her story about who they are, have been, and will become. This provides the unity and coherence to one's view of oneself and one's life (McAdams, 2009). Narrative is, then, how the individual construes the story of the self, the story of identity.²¹

In one of the most careful studies of moral exemplars, Walker et al. (2007) did extensive interviews with those who had been nominated in Canada for national prizes for heroic rescue or for sustained social service. They also interviewed a matched sample of individuals. They coded the interviews for how they structured their life narratives (McAdams et al., 2001). The moral exemplars did not differ very much in terms of Big Five trait characteristics or even lower-level characteristic adaptations, but they did differ markedly on how they told the stories of their lives. In comparison to the non-exemplars, their stories emphasized early secure attachment to parental figures, lack of early enemies, and presence of early support. Their narratives had more positive affective tone, they spoke more positively about communication, and were more likely to emphasize the needs of others. Finally, the stories they told also tended to see good coming out of bad occurrences, something McAdams calls a "redemptive" theme (McAdams et al., 2001). These strong differences at the narrative level suggest that this aspect or organization of identity is a central aspect of the moral self. We will see much more of this when we turn to Chapter 5.

4.3 Discussion

4.3.1 Conclusion

1. *Virtues are not what make us virtuous.*

Unlike Socrates and Protagoras, virtues do not seem to be useful constructs to use in psychological models of moral action. Ironically, virtue may not be a central psychologically active construct in a virtue-oriented moral psychology. Or, in the language we have used in this chapter, virtue may not be among the "springs of action." Virtue, like the Big Five traits, may describe the characteristic ways that individuals differ in acting, across a variety of situations (that is, at the nomothetic

²¹ For a broader discussion of this aspect of moral identity see Chapters 5 and 9.

level).²² But virtue names (honesty, courage, compassion) may not identify the psychological springs of action that organize and motivate each person's behavior. One needs instead to look at the fine-grained appraisal patterns, personal commitments, values, skills, life projects, and narrative identity across the lifespan to discover how people come to act honestly, courageously, or compassionately in certain domains. It will be these springs of action that generate the very human (in)consistency we show in living.

2. *There is no singular moral personality.*

Moral exemplars and saints (Colby & Damon, 1992; Flanagan, 1991), people who act in accordance with long and sustained commitments to moral goals, are only similar to each other at a very high level of abstraction (e.g. their life narratives often share a similar structure of optimism though they share no personality traits). The particular moral commitments they have and the characteristic ways each person has of achieving those commitments vary widely (Hill & Roberts, 2010). Thus, the empirical evidence for moral pluralism, even among the highly moral, is undeniable, at the levels of both characteristic adaptations (e.g. values) and narrative identity. Exemplars and saints are similar in that they are committed to the particular goods they are seeking. They vary widely in the moral goals they seek and in the ways they seek those goals.

4.3.2 Application

What does a person like me do in a situation like this? In the Introduction, we presented a bar fight that was handled in an unusual way by the justice system. One point of restorative justice programs is to provide psychological leverage to change the personality of offenders so that they will be less likely to re-offend upon release. The perpetrator in the bar fight assaulted someone wearing a yarmulke, though he did so in the presence of his friend who was Jewish. He was therefore confused when he was charged with a hate crime. The court procedure allowed the assaulter, the victim, and others to come together, after preparation, and tell their different stories. The victim was able to ask the assaulter, "What if I had not been wearing a yarmulke? Would you still have assaulted me?" This question, in the formal situation, brought some new perspective to the offender, who admitted his prejudice. As a result, the victim asked for a lesser sentence and the judge altered that reduced sentence to allow him to keep his job (two concessions that might induce gratitude).

There are individuals who have dedicated their professional lives to the success of restorative justice programs. The setting of the restorative justice hearing is itself "a situation like this" that influences the thinking and behavior of the participants. And that behavior then influences the perceptions of the participants and their later

²²Or one might be able to conceive of virtue as some integration of the pattern of behavior across all three levels of personality. This might match with some version of Aristotle's description of virtuous action: "Any one can get angry – that is easy – or give or spend money; but to do this to the right person, to the right extent, at the right time, with the right motive, and in the right way, that is not for every one, nor is it easy" (Aristotle, 340 BCE/1941, Book II, 1109a27). Virtue as integration might serve well as a psychologically active construct.

decisions (and commitments to action). Thus, each person in the hearing may have been calling on a slightly different “person like me” that was made relevant by the setting of the hearing. Throughout the process, the offender is invited to consider what “a person like me should do in a situation like this.” And the decisions that result from it, at least in this case, seem tailored to reduce cynicism on the offender’s part. In addition, it is likely there was change in the “situation like this” at the bar, in an attempt to change its reputation as a setting for bar fights. In this description we can see the complex interplay of personality and situation that is tracked by the most current personality approaches. Enduring ways of interacting with the world are altered by, and alter, the world in a way that usually provides a sense of coherence and consistency to actors.

4.3.3 Open Questions

1. *Only some consistency is moral consistency.*

Surely a sense of coherence and consistency in character is important to us. Also, most people want to think of themselves as moral. But this does not imply that the central consistency and coherence for every person is about morality. It is important for a functioning society that almost everyone have some minimal commitment to morality. But there are many other themes than morality that individuals construct to organize their lives, including being a meaningful part of a community or family, achieving status or success, or simply having fun. In Chapter 5, we will see that people vary on the extent to which they place morality as central to their identity. It is only the very few for whom moral projects are the center of their identity.

2. *What is the relationship of morality to our life themes/projects/stories?*

Since morality is not so often the central theme of individuals’ lives, what then is the relationship of morality to those other life themes, and how does this influence moral action? It may be that, to the extent that a life theme is moralized (construed as a moral enterprise), moral consideration and action will be limited to the domains within that theme. But the relationship may be more complex than that, and there is very little empirical investigation of these interactions to help us decide. It might depend, for instance, on the dimension of morality that is integrated into the theme (e.g. loyal achievement of success vs. sacred achievement of success). Or it may have more complex interactions depending on other personality variables (e.g. attachment schemas). We know enough to ask the question, but not to give answers.

Another question in this vein is whether those who have morality deeply integrated into a life theme act differently than those who do not. Chapter 5 will suggest that this is so, but we do not yet know enough to give details of those differences or to say how they come about.

3. *How does integration of morality into the personality happen?*

The perpetrator in the bar fight is one example of an intervention that might help someone integrate moral considerations into their life themes. But this is only one window on a longer process of rehabilitation for the perpetrator and even the more general process may vary greatly from one individual to the next. Chapter 9 will address this question and provide a range of models and processes. And just the range of the models will suggest how much work there is yet to do.

4.4 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Beck and Jackson (2020). “Idiographic traits: A return to Allportian approaches to personality.” An introduction to recently developed statistical methods that allow us to track the individual and their personality profile across time.
- Fowers et al. (2020). “The emerging science of virtue.” A theoretical model of the virtuous personality that decomposes virtue into four major components: knowledge, behavior, emotion/motivation, and disposition.
- Hill and Roberts (2010). “Propositions for the study of moral personality development.” An application of principles of personality psychology to moral psychology.
- Nisbett and Ross (1980). *Human Inference: Strategies and Shortcoming of Social Judgment*. A classic early statement and overview of the primacy of situational influence on judgment and decision making.
- Penner and Orom (2010). “Enduring goodness: A person-by-situation perspective on prosocial behavior.” A useful review of the literature in the domain of pro-social behavior that integrates the interaction of the person with the situation.
- The following four models all share a view of personality as adaptive, pluralistic and integrative:
 - Cervone and Tripathi (2009). “The moral functioning of the person as a whole: On moral psychology and personality science.”
 - McAdams and Pals (2006). “A new Big Five: Fundamental principles for an integrative science of personality.”
 - Mischel (2004). “Toward an integrative science of the person.”
 - Roberts and Wood (2006). “Personality development in the context of the neo-socioanalytic model of personality.”

References

- Aristotle. (340 BCE/1941). *Nicomachean Ethics* (W. D. Ross, Trans.). In R. McKeon (Ed.), *The Basic Works of Aristotle* (pp. 927–1112). New York: Random House.
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO Model of Personality structure. *Personality and Social Psychology Review*, 11(2), 150–166. doi:10.1177/1088868306294907
- Atkins, R., Hart, D., & Donnelly, T. M. (2005). The association of childhood personality type with volunteering during adolescence. *Merrill-Palmer Quarterly*, 51(2), 145–162. doi:10.1353/mpq.2005.0008
- Beccari, C. (1907). Beatification and Canonization. In C. Herbermann (Ed.), *The Catholic Encyclopedia*. New York: Robert Appleton Company. Retrieved from <http://www.newadvent.org/cathen/02364b.htm>
- Beck, E., & Jackson, J. (2020). Idiographic traits: A return to Allportian approaches to personality. *Current Directions in Psychological Science*, 29, 301–308. doi:10.1177/0963721420915860

- Bentham, J. (1817). *A Table of the Springs of Action*. London: Richard & Arthur Taylor.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110(2), 203–219.
- Browning, C. R. (1992). *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York: HarperCollins.
- Cawley III, M. J. (1997). The virtues scale: a psychological examination of the structure of virtue and the relationships between virtue, personality, moral development, and epistemological style. Unpublished doctoral dissertation, The Pennsylvania State University.
- Cawley III, M. J., Martin, J. E., & Johnson, J. A. (2000). A virtues approach to personality. *Personality and Individual Differences*, 28(5), 997–1013.
- Cervone, D. (2004). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, 56, 423–452.
- Cervone, D. (2021). Five paths to personality coherence: Integrative implications of the Knowledge-and-Appraisal Personality Architecture. *European Journal of Personality*, 36(3), 319–346. doi:10.1177/08902070211015599
- Cervone, D., & Bartoszek, G. (2013). Knowledge, appraisal and personality dynamics. In D. Cervone, M. Fajkowska, M. Eysenck et al. (Eds.), *Personality Dynamics: Meaning Construction, the Social World, and the Embodied Mind* (pp. 11–32). New York: Eliot Werner.
- Cervone, D., & Tripathi, R. (2009). The moral functioning of the person as a whole: On moral psychology and personality science. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 30–51). New York: Cambridge University Press.
- Colby, A., & Damon, W. (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. New York: Free Press.
- Colby, A., & Kohlberg, L. (1987). *The Measurement of Moral Judgment, Vol. 1: Theoretical Foundations and Research Validation; Vol. 2: Standard Issue Scoring Manual*. New York: Cambridge University Press.
- De Raad, B., & Van Oudenhoven, J. P. (2011). A psycholexical study of virtues in the Dutch language, and relations between virtues and personality. *European Journal of Personality*, 25(1), 43–52. doi:10.1002/per.777
- de St. Aubin, E., McAdams, D. P., & Kim, T.-C. (Eds.). (2004). *The Generative Society: Caring for Future Generations*. Washington, DC: American Psychological Association.
- De Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal misbehavior: scientists talk about the ethics of research. *Journal of Empirical Research on Human Research Ethics*, 1(1), 43–50.
- Dean, J. W., Brandes, P., & Dharwadkar, R. (1998). Organizational Cynicism. *Academy of Management Review*, 23(2), 341–352.
- Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Doris, J. M., & Stich, S. P. (2005). As a matter of fact: Empirical perspectives on ethics. *The Oxford Handbook of Contemporary Philosophy*, 114–152.
- Dovidio, J. F., Piliavin, J. A., Gaertner, S. L. et al. (1991). The arousal: Cost–reward model and the process of intervention: A review of the evidence. In M. S. Clark, *Prosocial Behavior: Review of Personality and Social Psychology* (vol. 1912, pp. 1986–1118). Thousand Oaks, CA: Sage Publications).
- Dovidio, J. F., Piliavin, J. A., Schroeder, D. A. et al., (2006). *The Social Psychology of Prosocial Behavior*. New York: Taylor & Francis.
- Erikson, E. (1963). *Childhood and Society* (2nd ed.). New York: Norton.

- Flanagan, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press.
- Fowers, B. J., Carroll, J. S., Leonhardt, N. D. et al., (2020). The emerging science of virtue. *Perspectives on Psychological Science*, 16(1), 118–147. doi:10.1177/1745691620924473
- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1), 5–20.
- Gollwitzer, P. M., Sheeran, P., Trotschel, R. et al., (2011). Self-regulation of priming effects on behavior. *Psychological Science*, 22(7), 901–907. doi:10.1177/0956797611411586
- Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2022). Small effects: The indispensable foundation for a cumulative psychological science. *Perspectives on Psychological Science*, 17(1), 205–215. doi:10.1177/1745691620984483
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029–1046.
- Graham, J., Nosek, B. A., Haidt, J. et al., (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101, 366–385.
- Gustafson, S. B., & Mumford, M. D. (1995). Personal style and person-environment fit: A pattern approach. *Journal of Vocational Behavior*, 46(2), 163–188. doi:10.1006/jvbe.1995.1011
- Haas, H. (2002). Extending the search for folk personality constructs: The dimensionality of the personality relevant proverb domain. *Journal of Personality and Social Psychology*, 82(4), 594–609.
- Haidt, J. (2007). Moral psychology and the misunderstanding of religion. Retrieved from http://www.edge.org/3rd_culture/haidt07/haidt07_index.html
- Haidt, J. (2008a). Morality. *Perspectives on Psychological Science*, 3(1), 65.
- Haidt, J. (2008b). What makes people vote republican? Retrieved from http://www.edge.org/3rd_culture/haidt08/haidt08_index.html
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that Liberals may not recognize. *Social Justice Research*, 20(1), 98–116.
- Haidt, J., Graham, J., & Joseph, C. (2009). Above and below left–right: Ideological narratives and moral foundations. *Psychological Inquiry*, 20, 110–119. doi:10.1080/10478400903028573
- Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, 31(1), 191–221.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.
- Hartshorne, H., & May, M. A. (1928). *Studies in the Nature of Character: Studies in Deceit*. New York: Macmillan.
- Hartshorne, H., May, M. A., & Maller, J. B. (1929). *Studies in the Nature of Character. II. Studies in Service and Self-control*. New York: Macmillan.
- Hartshorne, H., May, M. A., & Shuttleworth, F. K. (1930). *Studies in the Nature of Character. III. Studies in the Organization of Character*. New York: Macmillan.
- Helzer, E. G., Furr, R. M., Hawkins, A. et al., (2014). Agreement on the perception of moral character. *Personality and Social Psychology Bulletin*, 40(12), 1698–1710. doi:http://dx.doi.org/10.1177/0146167214554957
- Hill, P. L., & Roberts, B. W. (2010). Propositions for the study of moral personality development. *Current Directions in Psychological Science*, 19(6), 380–383. doi:10.1177/0963721410389168
- Hirsh, J. B., DeYoung, C. G., Xiaowen, X. et al., (2010). Compassionate liberals and polite conservatives: Associations of agreeableness with political ideology and moral values. *Personality and Social Psychology Bulletin*, 36(5), 655–664. doi:10.1177/0146167210366854

- Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, 4(2), 162–176.
- Huff, C., & Barnard, L. K. (2009). Good computing: moral exemplars in the computing profession. *IEEE Technology and Society Magazine* (Fall), 47–54.
- John, O. P., & Srivastava, S. (1999). The big-five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (pp. 139–153). New York: Guilford Press.
- Kochanska, G., & Aksan, N. (2006). Children's Conscience and Self-Regulation. *Journal of Personality*, 74(6), 1587–1618. doi:10.1111/j.1467-6494.2006.00421.x
- Kochanska, G., Koenig, J. L., Barry, R. A. et al., (2010). Children's conscience during toddler and preschool years, moral self, and a competent, adaptive developmental trajectory. *Developmental Psychology*, 46(5), 1320–1332.
- Kochanska, G., & Murray, K. (2000). Mother-child mutually responsive orientation and conscience development: From toddler to early school age. *Child Development*, 71(2), 417–431.
- Kohlberg, L., & Mayer, R. (1972). Development as the aim of education. *Harvard Educational Review*, 42(4), 449–496.
- Lapsley, D. K., & Lasky, B. (2001). Prototypic moral character. *Identity: An International Journal of Theory and Research*, 1(4), 345–363.
- Lapsley, D. K., & Narvaez, D. (2006). Character education. In W. Damon, R. M. Lerner, A. K. Renninger et al. (Eds.), *Handbook of Child Psychology: Child Psychology in Practice* (vol. 4, pp. 248–296). New York: Wiley.
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819–834.
- Little, B. R. (1999). Personality and motivation: Personal action and the conative evolution. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (vol.2, pp. 501–524). New York: Guilford Press.
- Matsuba, M. K., Hart, D., & Atkins, R. (2007). Psychological and social-structural influences on commitment to volunteering. *Journal of Research in Personality*, 41(4), 889–907.
- Matsuba, M. K., & Walker, L. J. (2004). Extraordinary moral commitment: Young adults involved in social organizations. *Journal of Personality*, 72(2), 413–436.
- Matthews, G. (2008). Personality and information processing: A cognitive-adaptive theory. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE Handbook of Personality Theory and Assessment* (pp. 56–79). Thousand Oaks, CA: Sage.
- McAdams, D. P. (2009). The moral personality. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 11–29). New York: Cambridge University Press.
- McAdams, D. P., & Pals, J. L. (2006). A new big five: Fundamental principles for an integrative science of personality. *American Psychologist*, 61(3), 204–217.
- McAdams, D. P., Reynolds, J., Lewis, M. et al., (2001). When bad things turn good and good things turn bad: Sequences of redemption and contamination in life narrative and their relation to psychosocial adaptation in midlife adults and in students. *Personality and Social Psychology Bulletin*, 27(4), 474–485. doi:10.1177/0146167201274008
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509–516.
- McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (pp. 139–153). New York: Guilford Press.
- McCrae, R. R., Costa, P. T., del Pilar, G. H. et al., (1998). Cross-cultural assessment of the five-factor model: The revised NEO personality inventory. *Journal of Cross-Cultural Psychology*, 29, 171–188.

- McGregor, I., & Little, B. R. (1998). Personal projects, happiness, and meaning: On doing well and being yourself. *Journal of Personality and Social Psychology*, 74, 494–512.
- Midlarsky, E., Fagin Jones, S., & Corley, R. P. (2005). Personality correlates of heroic rescue during the holocaust. *Journal of Personality*, 73(4), 907–934. doi:10.1111/j.1467-6494.2005.00333.x
- Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371–378.
- Milgram, S. (1974). *Obedience to Authority: An Experimental View*. New York: Harper and Row.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.
- Mumford, M. D., Murphy, S. T., Connelly, S. et al., (2007). Environmental influences on ethical decision making: Climate and environmental predictors of research integrity. *Ethics & Behavior*, 17(4), 30. doi:10.1080/10508420701519510
- Narvaez, D. (2010). Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Nisbett, R. E., & Ross, L. (1980). *Human Inference: Strategies and Shortcoming of Social Judgment*. Mahwah, NJ: Prentice Hall.
- Nussbaum, M. C. (1996). *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, NJ: Princeton University Press.
- Oliner, S. P., & Oliner, P. M. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, 57(1), 401–421. doi:10.1146/annurev.psych.57.102904.190127
- Penner, L. A., Dovidio, J. F., Piliavin, J. A. et al., (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, 56(1), 365–392.[AU: please check inserted vol/issue and updated date]
- Penner, L. A., & Orom, H. (2010). Enduring goodness: A person-by-situation perspective on prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature* (pp. 55–72). Washington, DC: American Psychological Association.
- Peterson, C., & Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. New York: Oxford University Press.
- Reimer, K. S., DeWitt Goudelock, B. M., & Walker, L. J. (2009). Developing conceptions of moral maturity: Traits and identity in adolescent personality. *The Journal of Positive Psychology*, 4(5), 372–388. doi:10.1080/17439760902992431
- Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of Personality Development* (pp. 11–39). Mahwah, NJ: Lawrence Erlbaum Associates.
- Roccas, S., Sagiv, L., Schwartz, S. H. et al., (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin*, 28(6), 789–801.
- Saucier, G. (2003). An alternative multi-language structure for personality attributes. *European Journal of Personality*, 17(3), 179–205. doi:10.1002/per.489
- Schroeder, D. A., & Graziano, W. G. (2015). The field of prosocial behavior: An introduction and overview. In D. A. Schroeder & W. G. Graziano (Eds.), *Oxford Handbook of Prosocial Behavior* (pp. 3–34). Oxford: Oxford University Press.
- Schwartz, S. H. (2006). Basic human values: Theory, measurement and applications. *Revue française de sociologie*, 4(47), 929–968.

- Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature* (pp. 221–241). Washington, DC: American Psychological Association.
- Shoda, Y., Lee-Tiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and consistency from intra- and inter-personal interactions. *Personality and Social Psychology Review*, 6, 316–325.
- Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology*, 67(4), 674–687.
- Snow, N. E. (2010). *Virtue as Social Intelligence: An Empirically Grounded Theory*. New York: Routledge.
- Snyder, M., & Clary, E. G. (2000). The functional approach to volunteerism. In G. R. Maio & J. M. Olson (Eds.), *Why We Evaluate: Functions of Attitudes* (pp. 365–393). London: Psychology Press.
- Snyder, M., & Clary, E. G. (2004). Volunteerism and the generative society. In E. de St. Aubin, D. P. McAdams, & T.-C. Kim (Eds.), *The Generative Society: Caring for Future Generations* (pp. 221–237). American Psychological Association.
- Thoroughgood, C. N., Sawyer, K. B., Padilla, A. et al., (2016). Destructive leadership: A critique of leader-centric perspectives and toward a more holistic definition. *Journal of Business Ethics*, 151(3), 627–649. doi:10.1007/s10551-016-3257-9
- Tomasello, M., & Vaish, A. (2012). Origins of human cooperation and morality. *Annual Review of Psychology*, 64, 231–255. doi:10.1146/annurev-psych-113011-143812
- Treviño, L. K., Weaver, G. R., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management*, 32(6), 951–990.
- Walker, L. J., & Frimer, J. A. (2007). Moral personality of brave and caring exemplars. *Journal of Personality and Social Psychology*, 93(5), 845.
- Walker, L. J., Frimer, J. A., & Dunlop, W. L. (2010). Varieties of moral personality: Beyond the banality of heroism. *Journal of Personality*, 78(3), 907–942. doi:10.1111/j.1467-6494.2010.00637.x
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, 86(4), 629–647. doi:10.1037/0022-3514.86.4.629
- Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review*, 8(3), 281–307. doi:10.1207/s15327957pspr0803_4
- Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, 21(3), 253–277. doi:10.1177/1088868316652279

Moral Identity and the Self

I want to do what is good, but I cannot do it: I do not do the good thing I want to do, but instead do the wrong thing that I do not want to do.

(Apostle Paul, Romans 7: 18 and 19)

Most who have tried to live up to some moral standard and failed will appreciate Paul's description of moral failure. In psychology, the less poignant term "judgment-action gap" has been used to describe this state of affairs. In Chapter 4, we have investigated this gap between desire and action. To understand it, we looked beyond simple behavior descriptions (honest, loyal, compassionate) to the crucial question: "What does a person like me do in a situation like this?" (Weber, Kopelman, & Messick, 2004, p. 281). This, of course, presumes two questions: "Who is a person like me?"; and "What is the situation like?" This chapter is about the first of those questions. We will find answers to the second question in Chapters 3, 4, 7, 8, and 9.

"Who is a person like me?" is a question in the domain of research on the self and identity. Moral identity, the extent to which moral commitments are integrated into the self, has been viewed as a crucial answer to the long-standing puzzle of the judgment-action gap, the problem that moral judgment often does not produce moral action.¹ In this chapter we will ask (1) What is the judgment-action gap and why does it appear? And (2) How does moral identity help us to understand that gap and to bridge it?

To do this, we will first need some understanding of (3) how the self is structured. We will find that there is not a single, unitary self, and no single way that morality might be integrated into it. This complex multidimensionality of the self and moral identity produces (4) an explosion of complexity in moral self-reflection, the conscious evaluation of the self in order to maintain consistency or moral integrity.² Self-reflection is easily susceptible to self-deception but also easily susceptible to

¹This puzzle is one of the reasons that this book is called *Taking Moral Action* rather than "Making Moral Judgments."

²In Chapter 9, we call these processes "evaluative moral formation" in order to emphasize their connection with the process of moral formation.

constructive critique by the self and others. We conclude (5) that a modest commitment to some values or norms³ in self-reflection within this complex landscape and modestly skilled objective evaluations of moral performance are together likely sufficient to produce some integrity.

5.1 From the Judgment–Action Gap to Moral Identity

The revival of research on the self and moral psychology was fueled by the widespread recognition that even when people are capable of making careful and mature moral judgments, this ability does not often result in moral action. The idea is at least as old as the writings of the Apostle Paul. Simply knowing the right/good/moral thing to do is not the same as wanting to do it, which is again different from a final translation into action.⁴ Hartshorne and May (1928) were among the first psychologists to empirically document this. With the children they tested, honesty in one situation had little to no predictive value for honesty in dissimilar situations. This pattern (or lack of a pattern) held for many other supposedly character-like attributes they tested. In 1968, Mischel reviewed the cumulative failure of psychologists to find cross-situational consistency for personality characteristics and famously proposed that there *were no simple personality traits* that were broadly predictive of behavior. A decade later, Blasi (1980) cataloged the consistently small correlations⁵ that were obtained between increases in Kohlberg’s moral stages and taking corresponding moral action.⁶ He labeled this disappointing state of affairs the *judgment–action gap*,⁷ and proposed *moral identity* as the motivational factor needed to bridge it. For Blasi, *moral identity* is the possession of a particular kind of self-concept (Blasi, 1984), a self-concept with high, even central, regard for issues of moral obligation, responsibility, and the good. We will analyze moral identity as a multidimensional aspect of the self-concept and follow the implications of this dimensionality for the judgment–action gap.⁸

The judgment–action gap is most problematic for those approaches, like Kohlberg’s, that have conscious moral judgment as the central aspect of the model (Frimer & Walker, 2008). These models depend upon psychological consistency pressures as the main engine to bridge the gap between knowing and doing. If one judges something to be the moral thing to do, it is uncomfortable to think that one has not done it.⁹ For Kohlberg, this uncomfortable consistency pressure comes from holding a

³How this commitment comes about is the subject of Chapter 9.

⁴The relationship among these three things is part of the philosophical puzzle called “weakness of the will.” See Stroud (2019) for a review of the long philosophical history of this problem.

⁵Both words in the phrase “consistently small” are important to note. There is a *very consistent correlation* (about 0.33) between moral cognition and moral action, but this consistency is at best “somewhat modest” (Walker, 2004, p. 2).

⁶See Chapter 7 for a presentation and critique of Kohlberg’s approach.

⁷Note that this formulation presumes the judgment comes (or ought to come) before the action.

⁸Chapter 9 is where we will explore how a person comes to have a strong moral identity.

⁹See Heider (1958) and Festinger (1957) for early versions of this motivational urging that is associated with thinking inconsistent thoughts.

principle (e.g. justice) and is thus about consistency with the principle.¹⁰ Blasi's moral identity (1980) approach is still centered on judgment, and so it too is powered by cognitive consistency. But it is consistent with the moral self (one aspect of the multidimensional self that has moral goals, etc.). Thus, for Kohlberg, immoral action is betrayal of a principle, but for Blasi, it is betrayal of the self, with presumably more motivating power (see Lapsley, 2008, for a review).¹¹

The judgment–action gap is more easily bridged in models of the self that actively include emotions and automatic evaluations as central players (Bandura, 1986; Mischel, 2004; Narvaez, 2010b).¹² Well-practiced emotions lead us to righteous anger or pity, and contain within them a motivation to act (see Chapter 8 for this). These approaches bridge the gap naturally (at least in theory) because they contain an action impulse as an integral part of the automatic response. Automatic, non-conscious evaluations (Haidt & Ksebir, 2010; Narvaez, 2010a) and emotions are combinations of cognitive judgment, self- and other-relevant emotion, and action orientation (Haidt, 2003; Lazarus, 1991b). The massive interconnectivity of emotional and judgment areas in the brain (Pessoa, 2008) attest to this blending of cognition, emotion, and action, and Moll et al. (2006) refer to (and also track over time) “cognitive-emotional” complexes in moral judgments. The cognitive appraisals in these models include reference to the implications of the situation for the self and its goals (Lazarus, 1991a, 1991b). Thus, these automatic/emotional approaches also have the self as a core component – not a meager, cognitive self but a richly complex one that includes motivational commitments and automatic tendencies toward action.

5.2 The Varieties of Self: Or What is Moral Identity?

In this section, we will consider the variety of psychological models of the self, considered under three propositions: that the self is multidimensional, variable over time and context, and often domain specific. In the process, these general characteristics of the self will be shown to be true also of the *moral self* or *moral identity*, which we define as those aspects of the self that are relevant to moral action. As *moral identity*

¹⁰Kohlberg founded his psychological system in part on Kantian models of reason as the source of moral judgment (Frimer & Walker, 2008; Lapsley & Narvaez, 2005). However, he mistakes the complexity of Kant's moral psychology when he claims reason alone should be the source of moral action. In a well-known footnote in Kant's (1785/2011, pp. 30–31) *Groundwork for the Metaphysics of Morals*, Kant bridges the gap between thought and feeling/motivation by treating the feeling of *respect* or *reverence* (Achtung) for a principle as both a cognitive *and* a motivational construct, so that the internal feeling comes not as outside influence but “through a motivation *self-wrought* by a rational concept” (durch einen Vernunftbegriff *selbstgewirktes* Gefühl). Note that “Vernunftbegriff” would be literally translated as “term of reason.” In contrast, Kohlberg sees all emotion as bias and ignores Kant's subtle moral psychology in favor of a simple rational actor model.

¹¹This self-betrayal is analyzed by Sartre (1943/1984) under the label of acting in bad faith.

¹²There are models of automatic moral cognition that do not include reference to the self, motivation, or emotion (Sunstein, 2005) and these have a similar judgment–action gap to bridge.

is used in much of the literature in moral psychology, this is the term we will use interchangeably with the *moral self* or *moral self-concept*.¹³

5.2.1 The Self is Multidimensional

Before the dawn of academic psychology, the Danish theologian Søren Kierkegaard pioneered the phenomenological study of the self. He portrayed the self as a threefold relationship that holds together the physiological, psychological, and theological/normative aspects of a person's life. Thus, the self is not a thing, or some shadowy substance, or a possession one might have. Instead, it is what one does, an ongoing narrative with which one stitches together past and future, being and becoming, conditions and possibilities. Such a self is seen as an agent of one's becoming, navigating a continuous and continuously revised set of relations, some imposed by the constraints and tensions of being human, others freely chosen as a response to those tensions, and all embedded within, or resting on, a narrative about ultimate meaning (Davenport, 2013; Garff, 2013).

In recent personality and social psychology literature, the self has been understood using the more limited metaphor of a cognitive network. This self-system is a multidimensional network that includes cognitive representations and evaluations of the actual, possible, ideal, future, etc. self (Cervone, 2004; Heatherton, 2011; Leary, 2007; Mischel, 2004). Critics of cognitive network approaches have complained that this cognitive bias in models of the self makes them inadequate for understanding our deeply motivating, complex, and evolving self-representations of who we are (Nucci, 2000, 2004). At one level, this complaint is telling. The rich experience of deeply meaningful activity (Colby & Damon, 1992), of social threats to our self-esteem (DeWall, 2009; Heatherton, 2011), and of personal crises of meaning (Koole, Greenberg, & Pyszczynski, 2006) cannot easily be modeled by a simple network of nodes linking traits that we think apply to us (Aquino & Reed, 2002).

However, most approaches do not see self-networks as such spare furniture in our psyche. Many approaches link our self-systems to self-reflections of competence and esteem (Bandura, 1986; Bandura & Locke, 2003; Heatherton, 2011; Leary, 2007) and thus implicitly or explicitly evoke the emotional commitments that come with a sense of self. Many systems explicitly include emotion in self-systems: Mischel and colleagues (Mischel, 2004; Shoda, LeeTiernan, & Mischel, 2002; Shoda, Mischel, & Wright, 1994) refer to the "cognitive-affective" processing system and Cervone (2004) explores the linkage of knowledge and emotional appraisal in judgments about the self. Other approaches to the moral self see aspects of skilled action (e.g. offering specific kinds of help to others) as highly automatized, scripted, and central to the self (Lapsley & Hill, 2008). Still other approaches see the themes that emerge from public and private narratives of the self as defining of the *moral self* (Colby & Damon, 1992; Huff, Barnard, & Frey, 2008a, 2008b; McAdams, 2009).

All models agree that the networks that make up our moral selves are multidimensional and include cultural and social influence, cognition, emotion, and intuition. And because they are action-oriented, they include highly practiced skill

¹³One can split many hairs over the differences between these terms. We do not have the space to do so; and we think avoiding it is worth any minor conceptual confusion it might create.

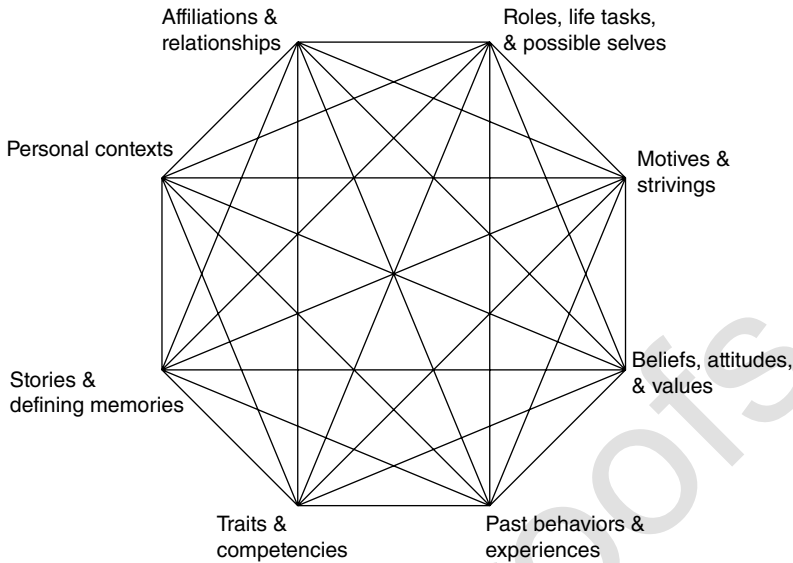


Figure 5.1 A model of some of the components of the self.

Source: Adapted from McGregor & Little (1998).

(McConnell, 2010). It follows that there will be a variety of ways that ideals, values, morality, and moral goals will be integrated into the self.

Figure 5.1 is a good example of the variety of ways that self can be constructed. It proposes a set of ways that themes, including moral themes, may be integrated into the self-concept (to answer the question “What is a person like me?”).¹⁴ This “diverse array” (McGregor & Little, 1998, p. 496) of elements includes at least sixteen different types of items ranging from memories of the past to hopes (and fears) for the future, from skills to attitudes to social roles. Items, of course, overlap and it is unlikely that the list is complete. But each of them suggests ways that morally relevant influences might be integrated into the self. For any single individual, a moral commitment to integrity might be experienced as, and be represented in, multiple things. For example, a computer scientist that Huff and Barnard (2009) interviewed in a study on moral expertise spoke of his architect father as an important role model. He noted that his father’s moral integrity was “incredibly influential” in his life.¹⁵ As interviewers we were confronted with a host of questions: How shall we talk about or represent this important aspect of this person’s life? We could conceptualize this as a *value* of

¹⁴We are using “moral identity,” “moral self-concept,” and “moral self” as synonyms in this chapter. These are aspects of the larger self. Indeed, when people are asked to rate what aspects of the self would most drastically change identity, the moral self appears to be the most important part of the self, surpassing even autobiographical memory (Strohmingier & Nichols, 2014).

¹⁵This is from a study of moral exemplars in computing, individuals who were widely known for their commitment to moral causes in the field of computing, selected by a panel of computer scientists and philosophers.

“integrity,” or as a motivating *story* about his father’s struggles in a system that was sometimes unfair? Or as a boyhood *memory* of Saturdays at the office with his father, or as a *possible self* that is like his father, or as the *skill/virtue* of courage he practiced in emulation of his father? Or all these together? Or of some at one time, and others at other times? There is evidence in the interview for all of these.

Each of these conceptions would likely tap something about the importance of his father, about the role his father played in his life and his particular moral self-concept. Some aspects might be more relevant than others, depending on the idiosyncratic way the exemplar thought about his father’s influence. And the best way to conceive of the father’s influence for this exemplar may not work at all for some other person with some other influence from some other family member. This is a theme we visited regarding trait and virtue theory in Chapter 4. Conceptions of moral identity or other variables that work well across a range of people are tapping *something* similar and work to predict behavior because of it. But they are likely not describing the particular way each individual experiences the world or thinks and feels about moral issues.

Most of those who do work in moral identity recognize this richness, and see their conceptualizations and measurement instruments as useful tools to tap something about moral identity, but not as definitive of all the ways moral identity operates or is experienced by the individual (e.g. Aquino & Reed, 2002, p. 1424; Narvaez et al., 2006, p. 982).¹⁶

5.2.2 The Self is Variable Over Time

There is great complexity within the web of self-relevant memories, descriptions, scripts, roles, etc. of how we represent our selves to ourselves. This multiplicity of self-identity is likely gradually developed over childhood, increasing in complexity as the child’s social world expands (Amiot et al., 2007). In this process, the self-concept becomes more and more multidimensional as its cognitive and emotional capacities to represent and interact with this complexity develop (Amiot et al., 2007). This differentiation of the (moral) self is influenced by, at least, parenting style (Hardy et al., 2010) and community context (Hart, 2005a; Hart & Carlo, 2005). As they develop, the multiple aspects of the self become more complex, varying in their interdependence, and in their overlapping attributes, narratives, themes, and emotional resonance (McConnell, 2010).

Coherence or consistency is often (though not always) imposed by the life narrative into which the individual retrospectively incorporates past, present, and possible future selves (McAdams, 2006). And these narratives themselves have implications for desired future behavior. But the expected consistencies may be limited. In our interviews of moral exemplars in computing we found some respondents insisted on integrating their family life into their stories, and others were perfectly willing to speak of their professional activity as though it was entirely independent of other aspects of their life (Huff & Barnard, 2009).

¹⁶See Frimer and Walker (2008) and Lapsley (2008) for reviews of how the moral self is measured. Though it does appear that each of the many measures is predictive of moral action, there is not yet enough in common among them to determine what is important and what is peripheral in these approaches and their concomitant measures. This is perhaps because the ways that morality can be integrated into the self-concept are legion.

5.2.3 The Self is Often Domain Specific

Situational influence can make some aspects of this web of meaning more relevant at one time than another. And though there are internal pressures for consistency, our sense of self is “highly differentiated and domain specific” (Lapsley, 2008, p. 46) and significantly influenced by our social surroundings (McConnell, 2010). Thus, the self we show up with in any situation may be slightly different depending upon the significance of that situation for us and upon the centrality of internal pressures for consistency in the moral domain, across situations. But we are not simply trapped by the situation: the influence is reciprocal. We often choose the situations we are in, in order to match our sense of self, and we can choose to influence those situations if there is some mismatch (Huff et al., 2008b; Killen & Dahl, 2021; Lapsley & Narvaez, 2004; McConnell, 2010).

One type of situational variation in the self that has received significant attention is variation based on social relations and community. We interact with many different individuals over time, and each sees a slightly different aspect of who we are (Andersen & Chen, 2002). Simply put, we show different selves, or different aspects of the self, to different others. We are, at various times, mother, teacher, friend, mentor, child, team member. Our interactions, or imagined interactions, with others can call out different interpersonal patterns of affect, belief, motivation, self-reflection, and self-regulation (Aquino, McFerran, & Laven, 2011; McConnell, 2010; Rusbult et al., 2009). This psychological grounding in schemas based on past experience in relationships is in fact what Bowlby and Ainsworth had originally proposed as the central mechanism in attachment (Bretherton, 1992).

However, we need not limit the dimensional variation in our selves to individual relations. We know that even in an individualistic culture, people vary in the extent to which they see their connections with others as significant aspects of the self and that seeing one's self as socially connected influences compassion and socially responsible behavior (Cojuharencu, Cornelissen, & Karelaia, 2016; Cross, Bacon, & Morris, 2000; Day & Impett, 2018; Vignoles, 2017). Snyder and Omoto (2008) note the importance of simple group membership in supporting volunteering, and how motivation for helping can change depending upon whether one is helping in-group or out-group members. For instance, in a series of studies, Stürmer and Snyder (2010) showed that individuals helped members of in-groups because of empathy but they helped members of out-groups because they were attractive or had some social value.

There is evidence that moral identity can widen one's “circle of moral regard” (Aquino et al., 2006; Reed & Aquino, 2003). This is reminiscent of the work of the Oliners (Oliner & Oliner, 1988) who showed that those who helped Jews and others escape the Holocaust had a much wider view of who “us” was, a view that included those being victimized. We show a different moral face to those who are “them” than we do to those who are “us.”

One might have great integrity about justice and care for those we include in the deserving community, while being callous and uncaring for those “other people.” Susan Opatow explores this dynamic under the label “moral inclusion/exclusion.” She has reviewed how moral exclusion is constructed in a moral ecology (Opatow, 2001) and how programs stressing moral inclusion can have their effects in reducing violence and injustice (Opatow, 2005). Moral *tightening* and *loosening* describes the ways these impulses can change in a culture over time in response to

challenges like scarcity or external threat (Gelfand, 2021). This moral exclusion is one of the crucial steps in the staircase to terrorism (Moghaddam, 2005; Moghaddam, Warren, & Love, 2014).

Our moral ecology interacts with our moral identity in other ways than the us-them dynamic of moral inclusion. Participation within a community actually helps to construct our sense of self. In a longitudinal study of high school students, Pratt et al. (2003) found that self-ideal (measured by self-report of traits one strives for, such as being a good citizen, fair, just, caring) predicted community involvement, but over time it was community involvement that led to increases of these self-ideal traits. Thus, active participation in community can produce a virtuous cycle of self-ascription of traits, followed by further community service. Hart (2005a, p. 260) argues: “If the notion of identity is to contribute to an understanding of moral functioning, then it must be a construct with deep roots in a social world.” These roots will need to be both deep and specific to each individual and his or her personal history and moral ecology. Piliavin and colleagues (Grube & Piliavin, 2000; Piliavin & Callero, 1991) have shown the importance of “specific role identity” (as, e.g. a blood donor or a cancer society volunteer) in supporting volunteer activity. Rule and Bebeau (2005) and Huff and Barnard (2009) have also tracked the importance of specific identity themes in professions.

So, the “deep social roots” of the self, and thus of the *moral* self, may well be quite specific, not easily susceptible to consistency pressures across domains, and may not reach across roles or domains to influence moral behavior in other contexts. In this way, the specificity and multidimensionality of the self can become a threat to pressures of consistency across domains; or what in moral language is often called integrity. To investigate this issue, we will need to ask how consistency pressures, or pressures toward integrity, operate in the multidimensional self.

5.3 How Does the Moral Self-Concept Influence Action?

Supposedly, moral identity influences moral action in that the more that moral ideals/values/models/narratives/etc. are integrated into the self-concept, the more individuals’ moral actions are motivated and structured by that moral identity. Chapter 9 traces this integration. We mention it here to remind us that these moral identity models are designed to solve the problem of the judgment–action gap discussed earlier.

One important thing to note is that the empirical gap is less broad than often thought. Blasi’s (1980) review and several subsequent ones (Frimer & Walker, 2008; Jennings, Mitchell, & Hannah, 2015; Lapsley, 2008; Lapsley & Narvaez, 2005) have established a moderate relationship between moral identity and action, particularly when the action is seen as self-relevant (Reynolds & Ceranic, 2007). This helps us to see the judgment–action gap as something that, from the perspective of the self, is only consequential when the judgment and the action conflict with each other in ways the self judges to be relevant.¹⁷

¹⁷One can, of course, disagree with someone about whether an action is relevant, and even question one’s own actions or integrity. This process makes judgments of relevance central to understanding integrity.

Still, the self that is supposed to drive moral action has been revealed to be multidimensional, variable, and domain specific. One can see this as fragmentation of the self (Gergen, 2000), which makes the gap wider, or at least deprives the actor of a strong incentive toward unity of judgment and action. When judgments of relevance can splinter along lines of domain, situation, and aspects of the self, there seems to be little in the model that pushes for unity or integrity. The unity of the moral self might thus dissolve into a postmodern mash of singular, conveniently changed, isolated hopes, urges, and defenses (Frimer & Walker, 2008, p. 344) with no central organizing principle to provide internal consistency or integrity.

Alternatively, one can see the multidimensionality as authentic adaptation to the complexity of life, in which aspects of the moral self cleave the social world based on its different moral ecologies. In this view, some variation in response to situational difference might be respectable and even expected (Shoda et al., 2002). Domain limitations on moral action may well be the result of what existentialists call the need to reconcile life's unavoidable contradictions. And it may be that a well-developed moral identity allows for thoughtful movement between moral ecologies, or the careful integration of values one might think are opposed (see the concept of integration in Section 5.3.1.1 and the concept of self-transcendence in Chapter 9). One might then think of moral identity as less a moral straitjacket and more a well-coordinated wardrobe of clothes. It may best be thought of as integrating the moral tensions we experience in the many domains of our life rather than simply serving as a bulwark against failure.

This review of the multidimensionality of the (moral) self should lead us to be careful in framing any demand for moral consistency, integrity, or unity. There are at least two issues here: when we speak of a drive toward consistency *what* items need to be consistent with each other, and *how* (in what way) are these items consistent?

5.3.1 Two Forms of Moral Integrity

What are the aspects of our self that we are longing to have consistent or to show integrity? Blasi (2005) provides a suggestion by outlining two kinds of self-integrity: *identity integrity* (internal coherence among ideals, principles, values, and other aspects of self-conception) and *behavioral integrity* (coherence of external behavior with internal conceptions).

5.3.1.1 Identity Integrity Identity integrity is the postmodern puzzle we have been reviewing. To what extent is the conception of self across domains and values integrated? We have seen that there is ample evidence that most individuals' self-conception varies across domains and with situations. There is evidence from Schwartz's (2016, 1987) cross-cultural work in values that there is even a fundamental incompatibility built into the structure of the way we think about values. In his well-validated circumplex model, values are arranged in a circle, with proximity on the circle representing how compatible they are in terms of the extent to which people report adherence to them. For instance, the values of benevolence and universalism lie next to each other, while power assertion and individual achievement lie next to each other on the opposite side of the circle (see Figure 3.1 in Chapter 3). This

opposition of values to each other suggests that integrity among one's values will be a difficult achievement. However, Walker, Frimer, and colleagues (Frimer & Walker, 2008; Frimer et al., 2011; Walker & Frimer, 2015) provide evidence that it is the integration of some of these supposedly opposing values that especially marks moral identity integrity and that also predicts moral behavior. In their *reconciliation model*, they identify people for whom *agency* (an individualistic value) and *communion* (a moral value of benevolence and universalism) co-occur. These peoples' agency is centered on communion with others, thus reconciling the supposed tension between them. Their defining moral goals are *both agentic and centered on others*. Here is an example they provide from an interview (Frimer & Walker, 2009, p. 1678):

In any action we take, in any choice we make, we can have a positive or negative impact on the world ... In terms of good habits, treating all people with dignity and respect regardless of their situation in life or how similar or different they may be

This dedication of the (agentic) choosing self to (communal) others, provides a theme of unity in the moral self that can guide moral behavior. We review the evidence for this unity in Chapter 9, Section 9.3.1.1.

5.3.1.2 Behavioral Integrity: Or How to Overcome the Judgment–Action Gap The judgment–action gap we have documented in this and other chapters is precisely the issue of behavioral integrity: the extent to which external behavior coheres with internal commitments (Blasi, 1980). We know this gap exists and can document some of the processes that narrow it. But where can we find an impetus toward more behavioral integrity? There needs to be some perceived inconsistency between one's behavior and something internal for there to be some motivation to reduce it. What is the internal aspect that serves as a reference point? We present here a variety of candidates, and each one may, under some circumstances, be the aspect that drives the perception of consistency or inconsistency in behavior:

- *Principle.* Consistency with a principle was proposed as the motivating factor by Kohlberg (1958, 1971). Whether this is one all-encompassing principle or a set of principles is unclear, even in later versions of the theory (Kohlberg, Levine, & Hewer, 1983). However, it reappears in our discussion in Chapter 9 of the idea of a higher goal or telos that reorganizes our values.¹⁸
- *Self-concept or identity.* Blasi's (1980, 1984, 2005) influential work suggests that what gives a perception of consistency-motivating force is when the behavior is seen as consistent or not with one's understanding of who one is, one's moral self-concept. We have reviewed this literature in this chapter.
- *Self-within-context.* Several authors have argued that a sense of responsibility or obligation does not arise out of a simple consistency pressure of behavior with self-concept, but from self-concept within social systems that constitute one's social

¹⁸We should note that this higher goal need not appear phenomenologically to the moral actor as a principle. It could show up in any of a variety of guises. In Section 5.2.1 we give an example of how the influence of a father might be integrated into the moral self in many ways.

and cultural relations (Haidt, 2001; Hart, Atkins, & Southerland, 2006; Hart & Matsuba, 2009; Markus & Kitayama, 2010; Power, 2004). Cultural meanings and social surround help to identify those aspects of the self that are the most relevant to determining behavioral integrity in any context. This is an expansive definition of the self, and one that challenges the simple dichotomy of person *vs.* situation.

- *Emotion.* An emotion has a natural action tendency (Arnold, 1960; Ellsworth, 2013), hence an action or a lack of action can be inconsistent with this action tendency.¹⁹ A variety of researchers have identified moral emotions as containing the kind of motivational power that can drive behavioral integrity (Haidt, 2001, 2003; Oliner & Oliner, 1988; Tangney, Stuewig, & Mashek, 2007). This is in part because the process of “having” an emotion involves a judgment of the relevance of some occurrence to the self. So, one can treat emotions as a separate source of behavioral integrity or as one process through which judgments of self-relevance have their effect.

There is empirical support for all four of these reference points to serve as the contrasting pole that produces pressures for behavioral consistency, or behavioral integrity. Rather than being alternative explanations, it may well be that they act independently or in concert. There is simply not enough theory or data to allow more than speculation about when or how each might have influence.

5.3.2 How Are These Aspects Consistent?

What do we mean when we say that some action is inconsistent with some standard? What kind of comparison for consistency are we making? In almost all the psychological literature, the default understanding of consistency has been logical consistency. Logical entailment has been used as the measure of moral consistency in part because it is most suited to the project of finding a logically consistent and unquestionable foundation for our sense of right and wrong.²⁰ In our later list of “ways of being consistent”, logical consistency stands out because it seems simultaneously to answer questions about how we make moral judgments *and* whether those judgments are ultimately justified. It was for this reason that Kohlberg (1958) adopted the logical consistency criterion (Lapsley & Narvaez, 2005).

None of the other kinds of consistency are “foundational” in this way. They are much more attempts to simply describe how people make consistency judgments, or “feel” that things are inconsistent. Though they cannot offer a logical connection to a principled position, they can still differentiate among better and worse or more and less in the amount of consistency. And so, they can still provide for some motivation to increase consistency.

¹⁹See Chapter 8 for more detail.

²⁰See Rudd (2007) on narrative consistency not being logical consistency or suitable for a foundationalist approach. See also Chapter 7 for Lawrence Kohlberg’s long, foundationalist battle against relativism.

- *Logical entailment.*²¹ Consistency defined as logical entailment seems to set up a dimension on which more is always better. Colby and Damon (1992) talk about their moral exemplars often striving to broaden and deepen their moral commitments over the life span. Frimer and Walker (2008, p. 349) suggest that the attainment of complete logical consistency of moral motives within and across all domains and action is “a practical impossibility.”²² It may even be paradoxical, not least because it can itself distract from the goal of striving for the good.
- *Moral ledger.* But there is a less demanding interpretation proposed by Nisan (1991) as a “moral ledger.” Nisan provides evidence that people trade off moral credits and debits and seem satisfied when the overall balance is above some positive threshold. This relaxed approach allows, of course, for its own kind of abuse in *moral licensing*, the belief that having done something good can then entitle one to be more lax in other areas (Conway & Peetz, 2012; Monin & Miller, 2001).
- *Emotional consistency.* Emotional experiences such as elevation (Algoe & Haidt, 2009; Haidt, 2000; Vianello, Galliani, & Haidt, 2010), empathy (Oliner & Oliner, 1988; Preston & de Waal, 2002; Tangney et al., 2007), and compassion (Goetz, Keltner, & Simon-Thomas, 2010; Oveis, Horberg, & Keltner, 2010; Stellar & Keltner, 2014) produce motivations to do good. We will explore the nature of these motivations in Chapter 8, but they can be signals that something is out of balance and needs to be corrected or is approaching balance and should be pursued. In addition, they can be cultivated: one can learn to be more (or less) emotionally sensitive (Weng et al., 2013; Wong, 2015) and can make this emotional development a goal (Koole, 2009; Koole & Veenstra, 2015; Kuhl & Koole, 2004).
- *Social consistency.* We have already reviewed the work on the self as embedded in culture and situation. Rather than simply viewing this as a corrupting influence (from the perspective of logical entailment) one can also see it as social influence toward consistency and integrity. Colby and Damon (1992) recount numerous instances in which it was social influence that led their exemplars into deeper commitment to causes, or to widen the horizon of their concern. The consistency pressures here can be thought of as mutual social influence in a way that greatly enhances the social aspect of Haidt’s (2001) social intuitionist model. By sharing their moral reactions, people mutually influence each other’s intuitions about the good.
- *Instrumental consistency.* Work on moral exemplars (Colby & Damon, 1992; Huff & Barnard, 2009; Plaisance, 2014) suggests that many view their commitments as moral projects, and evaluate their actions based, at least in part, on how they are compatible with achieving their moral goal. This has the problematic aspect that one might choose immoral methods to attain a moral goal (Hart, 2005b;

²¹We use “logical entailment” here to mean what philosophers Beall and Restall (2019) refer to as “the relation between premises and conclusions in valid arguments” and call “logical consequence.” Exactly what *this* means is, of course, a matter of dispute and Beall and Restall review that dispute.

²²This practical impossibility in the face of the demand for complete consistency is related to the sense of inadequacy in the face of the infinite ideal (e.g. the highest good) that has been puzzled over from ancient times to Kierkegaard and beyond. It is also found in Reinhold Niebuhr’s reflections on the “impossible possibility” (Niebuhr, 1935, chapter 7).

Skitka, 2010; Skitka & Mullen, 2002). Not all examples of this expediency are negative. The theologian Bonhoeffer's participation in a plot to assassinate Hitler and Nathaniel Borenstein's decision to consult for the military are two examples that are arguably positive. But it may nonetheless describe the nature of the motivational process behind many goal-driven moral actions. It seems quite similar to what we call in Chapter 9 self-transcendence, when the self transcends not only one's goals but also ethical convictions toward a higher calling in a certain place and time.

- *Narrative consistency.* Many theorists are now exploring the role of narrative in psychological functioning. Narratives have a structure and unity that is not univocal, in the way that logical entailment must be, but is instead “polyphonic” (McAdams, 2006, 2009), with multiple threads sometimes competing, sometimes cooperating. The telling of narratives serves a variety of purposes, some of them centrally moral (McAdams, 2009; McAdams & Pals, 2006; Tappan & Brown, 1989). All people tell narratives for social-bonding purposes, but people also tell stories to direct and guide their future behavior and to establish and maintain self-continuity (Bluck et al., 2005; De Silveira & Habermas, 2011; Habermas, 2011; Rasmussen & Habermas, 2011). These latter two purposes are both in part moral and can be thought of as relating to behavioral integrity and identity integrity respectively. Frimer and Walker (2009) have shown that it is just this kind of *narrative* unity between the opposing themes of benevolence/universalism and power assertion/achievement that predict successful moral action. Thus, narrative consistency can provide a sense of continuity, psychological pressures to consistency, and motivation to act, while being more flexible than simple logical entailment consistency.

We now have a quite complex picture of what it might mean to say one has integrity within one's moral identity by achieving consistency between some moral reference point and one's (im)moral behavior. Table 5.1 gives a picture of that complexity.²³ Rows show various kinds of reference points (e.g. a principle) and columns show how each might be compared with an action. There are four different kinds of reference points and a variety of modes of consistency.²⁴ The cells contain interpretations of what might psychologically be happening to produce consistency pressure in an actor within that cell.²⁵ For instance, where self-in-context and emotional consistency intersect, an actor might ask, “Is my emotion/motivation for this action consistent with my role as a doctor/daughter/feminist?” Or where principle and emotional consistency intersect, they might ask, “Is my emotion/motivation for this action consistent

²³Cervone (2021) provides a somewhat simpler overview of five ways one can construe personality coherence: (1) belief-based; (2) goal-based; (3) evaluative standards-based; (4) intra-psychic (interrelations among personality systems); and (5) phenomenological. These could be used to add even more (and perhaps fatal) complexity to this approach.

²⁴More or fewer might be imagined. See Cervone (2021) for a recent parallel proposal of different aspects of thinking about personality consistency.

²⁵We leave filling in the cells, and offering alternative interpretations of what a cell might mean, as an exercise for the reader. The rows and columns are not comprehensive and may need to be tailored to reflect the particular background of any individual.

with a principle of respect for others?” One can easily take advantage of this large array of possible self-reflections to find the most convenient comparison for one’s purposes, or one could take the time and resources to look for a more global kind of consistency.

Consistency among the rows in Table 5.1 might be called identity integrity. But bridging the gap, or behavioral consistency is what happens in each cell. A person who feels no pressure for consistency among the various kinds of judgments in the cells, or chooses them based on convenience, can be what the philosopher Frankfurt (1971, p. 11) calls a “wanton.”

How much consistency is enough? The philosopher Schwitzgebel (2019) makes a data-based case that something like a B+ is the criterion most well intentioned people have. However, there is little research on this and how it might be modified by culture, domain, moral identity, etc. It is likely that moral exemplars, at least in the relevant domain, have a higher standard. And it is possible also to dissociate some particular performance of, say, dishonesty, in a domain as a role-excused “free trait” that does not count against overall feelings of personal integrity (Schmidt & Poole, 2020).

Doing this well, of course, requires skills like self-reflection, emotional regulation, and moral imagination, which we cover in Chapter 6. It also requires appropriate, or at least a modest, commitment to some values or norms as discussed in Chapter 9, Section 9.3. And with the appropriate skills, norms, etc. one can reflectively assess the various consistencies in Table 5.1 to approach a *social-emotional reflective equilibrium*²⁶ that we describe in Chapter 8. Here lies the trap described at the beginning of the chapter by the Apostle Paul: there are many ways to fall off the path.

Many of these approaches to moral consistency come with a price: they open the process up to other motives (Cervone & Tripathi, 2009; Frimer & Walker, 2008) than the balanced consistency we describe as *social-emotional reflective equilibrium*. Many authors view this as a serious liability (Blasi, 1999, 2009; Kohlberg, 1963; Walker & Frimer, 2009). If consistency with moral principles is only one of many possible motives for moral action, then one is faced with diluting the drive to unity and the value of integrity, and we may be left with a fragmented moral sense. Whether the moral sense is – in fact and in lived experience – fragmented and cobbled together or carefully constructed as a polyphonic narrative is in the end an empirical question – but one with significant implications for judgments about morality.

5.3.3 If Moral Evaluation is So Complex, Can We Ever Hold Each Other Accountable?

A robust societal morality requires that we be able to hold each other responsible for our actions. The seductive call of logical consistency accounts is that they provide us with firm ground on which to make judgments of accountability. But one does not

²⁶One can easily imagine what *social-emotional reflective equilibrium* consists of by simply scanning the types of consistency in Table 5.1. We describe it in other chapters as a process for mature, or expert, moral action. This is what might count as a more global kind of consistency, integrating the table.

Table 5.1 The variety of moral consistencies.

Mode of Comparison: <i>In what way</i> is the action being compared with the reference point?							
Reference Point: <i>What</i> is being compared with the action?	Logical	Moral Ledger	Emotional	Social	Instrumental	Narrative	
	Principle/ Value	Action matches principle?	Emotion of action is based in principle?	Action matches values of others/roles?	Action achieves valued goal?		
	Moral Identity	Action changes positive balance of identity?	Action matches personal character based on others/roles?	Action matches professional identity?	Action fits in story of moral identity?		
	Self-in-Context	Action changes positive balance of identity in context?	Action matches important others/roles?	Action achieves moral goal shared with others or of role?	Action fits in story of self in context?		
	Emotion	Emotion of action matches preferred emotion?	Action matches current emotion?	Action achieves emotional goal?	Action fits with emotional tone of personal narrative?		

need strict logical consistency in interpersonal relationships to have influence. We can still hold each other accountable for our actions “so long as there exists an agentic, unifying process [that] attempt[s] to mend inconsistency and [has] some (incomplete) success” (Frimer & Walker, 2008, p. 346).²⁷ Each of the modes of comparison in Table 5.1 have ways of evaluating consistency, and all are subject to dispute. But they all share at least a modest impulse to consistency between the reference point and the action and some shared ground for recognizing and debating that impulse.

Where might this drive for unity come from? First, there is a kind of logic within each mode of evaluation. Second, we review in Chapter 3 ways that the social surround might support the internal logic in each mode. Third, Weinstein, Przybylski, and Ryan (2013) propose three integrative characteristics of self-regulation that support the unifying processes: (1) conscious access to one’s emotions, motivations and values; (2) taking responsibility for one’s emotions, decisions, and thoughts; and (3) non-defensiveness in response to challenge. None of these sound (or are) easy, since they often involve an honest response to failure. But they involve skills that can be learned (as we discuss in Chapter 6).

5.4 Critiques of the Moral Self

There is a clear case to be made for the importance of moral identity in moral action. But the concept is not without its critics. We have mentioned earlier the complaint of Nucci (2004) that many models of moral identity are too simplistic, and we have seen that there are multidimensional models that respond to this complaint (Amiot et al., 2007; Hart & Carlo, 2005; McConnell, 2010; Reed & Bolton, 2005). But in his wide-ranging critique of the moral self-construct, Nucci (2004) is also worried that some models of moral identity essentially reduce morality to a sort of moral egoism: “I do it because I want to.” Frimer and Walker (2008, p. 352) aptly respond that this is only a problem to the extent that the moral identity is conceived to be the *only* operating factor. There is evidence, for instance, of a variety of influences on volunteering, including wanting to have a particular experience or collect a line on one’s vita (Snyder & Omoto, 2008). But this is rarely the only influence, and even when it is, it might over time be transformed into other motivations (Hart & Matsuba, 2007). Once one includes moral reasoning, moral emotion, embeddedness in a moral ecology, or other influences, then we are back to a model of balancing influences. In addition, if moral exemplars’ moral identities are centered on helping others, one must accuse them of a particularly complex and cynical sort of egoism that includes helping others as a façade for mere self-enhancement (Badhwar, 1993). There seems little evidence to support this contorted critique.

²⁷It is this issue of whether our moral psychology allows us to insist on accountability or responsibility that is the central motivation behind Kohlberg’s and others’ insistence on a non-relativist moral psychology (Blasi, 2009; Kohlberg, 1958). The approach in this chapter certainly could be relativist, but the “modest drive to unity” is, we think, at least equally compatible with an approach called pluralism (Flanagan, Sarkissian, & Wong, 2008), which avoids the radical relativist abandonment of any standard.

A more problematic critique is that the role of individual moral identity may vary dramatically by culture (Cervone & Tripathi, 2009; Miller, Goyal, & Wice, 2017). The self is constituted differently in different cultures (Markus & Kitayama, 2010) and even at different levels of social status in the United States (Bowman, Kitayama, & Nisbett, 2009). Individuals in India, for instance, view not helping others as a moral failing regardless of the relationship with the other, while Americans see this as a moral issue only for close others (Miller, Bersoff, & Harwood, 1990) – a distinction that maps onto cultural differences in how embedded the self in each culture is in social networks (Markus & Kitayama, 2010). In reviewing three decades of research on the self and culture, Markus and Kitayama (2010, p. 421) note that “the study of culture and self has led to the realization that people and their sociocultural worlds are not separate from one another. Instead they require each other and complete one another... one cannot be a self by one’s self.” This makes it clear that the issue is not *whether* the self is involved in moral action in Indian or other collectivist cultures but *how* the self is differently construed and constructed, and *how* it contributes to moral action in different cultures. Thus, this criticism of the cultural inconsistency of moral identity effects can be recast as an invitation to explore the different ways that moral identity acts in different cultures, depending on how the self is construed in those cultures.

Another critique of the self as a key aspect of moral action comes from work by those who claim the self’s primary function in the psyche is to “facilitate people’s social interactions and relationships” (Leary, 2007, p. 317).²⁸ This might be viewed as a “selfish” motivation of self-enhancement or social accommodation: The self experiences “moral” emotions and acts on them merely to achieve social status or success. But this is a narrow view of Leary’s sociometer approach and it does not take into account the seemingly other-centered focus of moral emotions (Aktipis & Kurzban, 2004). It also confuses two levels of explanation, the ultimate (evolutionary) and the proximate (motivational) (Preston & de Waal, 2002) – it is one thing to say that the self supports moral behavior “to facilitate relationships” at an evolutionary level, and quite another to claim that the experience of the motivation to be moral is primarily based in the desire to have better relationships.

A final critique stems from the cross-cultural finding that traditional approaches to spiritual growth almost uniformly speak of denying the self, or the death of, or non-existence of the self as a necessary (and continual) aspect of becoming a whole, or virtuous, person (see Augustine, 1986 for an analysis in Christian, Buddhist, and Hindu monasticism). The Christian monastic Merton (2009) suggests that moral formation has to do with a life-long *transformation* of the self, which does not leave the self fully behind but rather transforms it from a mere egocentric self to what he calls the “true self.” This claim has a long religious, and specifically monastic, history (Merton, 2009) and can be found across many religious expressions (Shulman & Stroumsa, 2002). On its face, this emphasis on self-denial seems a direct contradiction to the sense of self as the motivator of action that we have found in this chapter.

²⁸Note that in Table 5.1, there is a column that is primarily driven by consistency pressures from important others and social roles. The table thus incorporates this critique as a kind of consistency.

If, as monastic language would have it, one must become detached from the self, how can this motivate moral action? There are two clues in psychological theory that may help us here. First, among his later theoretical formulations of the hierarchy of needs, Maslow differentiated the highest level of *self-actualization* into two parts: self-actualization and *self-transcendence*. Self-transcendence involves “seek[ing] a benefit beyond the purely personal and seek[ing] communion with the transcendent ... com[ing] to identify with something greater than the purely individual self, often engaging in service to others” (Koltko-Rivera, 2006, p. 306). Thus, in self-transcendence, one seeks something beyond self-interest, and this becomes more important than, or possibly identified with, self-interest. One might find this happening in several of the examples with which we began this book: The Munsons moved from taking in disabled foster children to founding a social service facility for disabled children; the Mothers of the Plaza de Mayo moving from concern for their child to larger social and political goals; and Thomas Merton seeking to transform the self in search of God and thereby becoming a peace advocate.²⁹

A second approach that shares the pattern of de-emphasizing the self comes from the research program of Walker and Frimer (2015). They and their colleagues have done extensive work on moral exemplars using both quantitative and qualitative methods. They describe their moral exemplars as having connected the good of others, or of some ideal, to the central moral concern of self and view this as an integration of two values, with one (self-agency) becoming an instrumental path to achieving the other (communion).³⁰

We should be careful not to assume a harmony among these approaches. The monastic approach emphasizes more detachment from the self than do the other two models, and it includes vast variation among different monastic traditions in different religions as well as non-religious forms of spirituality. Maslow seems to find self-transcendence attaching to a wide range of transcendent objects, as opposed to the other two approaches. And Walker and Frimer (2015) seem to view the process as a natural developmental path for most people, while for Maslow and the monastic approach this achievement seems more exceptional, in some traditions only truly achieved as a rare form of enlightenment. But all agree that there is a path to moral action that involves some form of decentering or transforming the self. We will explore this path in detail in Chapter 9.

5.5 Discussion

5.5.1 Conclusion

1. *The integration of morality into the self is important in moral action.*

Many research programs and theoretical models point to the importance of moral identity, the integration of morality into the self, as serving an important role in

²⁹There is a dark side to self-transcendence (Koltko-Rivera, 2006; Skitka & Mullen, 2002) so that simple self-transcendence cannot serve as a marker of morality per se.

³⁰Again, see Chapter 9 for an introduction to the Walker and Frimer (2015) research program and a reprise of this discussion.

the springs of moral action. There are regular, moderate-to-strong effects of moral identity in shaping moral behavior, particularly when one matches the situation and behavior to the specific commitments of moral identity. The importance of matching the situation and behavior to specific commitments is parallel in structure to the patterns we found in Chapter 4. Indeed, they are likely the same findings, as the measures used in personality are probably tapping aspects of moral identity. There is, then, mounting and converging evidence that moral identity plays a crucial part in the processes that initiate, regulate, and maintain (or derail, degrade, and misdirect) moral action. Moral identity is not, of course, *all* that influences moral action, as the other chapters in this text will attest. There are likely kinds of moral action that are somewhat unrelated to moral identity (e.g. habitual action – see Chapter 6). But moral identity is a prime candidate for playing a central role in much moral action.

2. *The moral self is multidimensional and variable over time and domains.*

The multidimensionality of the moral self suggests that there is no single “bridge” across a single “gap.” As the table of moral consistencies (Table 5.1) shows, there are multiple reference points (rows) against which one might judge a moral action, and then multiple modes (columns) in which this action might be interpreted in light of the reference point. The bridge and gap metaphor portrays an isolated individual vainly straining to overcome a gap between ideals and action, a gap that should not be there but constantly haunts the reflective actor (as the Apostle Paul was haunted). But it is far more complicated than this, and perhaps less dire. There is not one single value but a pluralism of values. And the values may be represented in a variety of ways in the complexity of the self and its representations of the moral ecology. Different aspects of the self become relevant as one navigates the moral ecology, and the self develops and changes over time as a result of these interactions and the ways they are evaluated. Finally, self-reflection, the skill of evaluating the fit of one’s action to one’s moral identity is itself legion in its modes. In this complex landscape, one should not be haunted by gaps but instead encouraged by woven networks of consistency that endure across the complexities of the moral ecology, of the self, and of self-reflection (see Chapters 6 and 9).

3. *Self-reflection as a guide for moral action is very flexible. From the standpoint of behavioral integrity, this is both a feature and a bug.*

The multitude of reference points and modes of comparison make the self-reflection process very flexible: a positive feature given the complexity of our moral ecologies and of how we think of our selves. It provides for coherence balanced by flexibility in how we adapt our selves to the complexities of human life. However, one can see this flexibility as providing too many opportunities for self-serving interpretations of one’s actions. Chapter 7 provides an overview of some of the ways we use reason to construct consistencies that allow us to excuse or support immorality. What then supports the impetus for consistency in moral action that comes from self-reflection? We argue in this text that many influences (moral ecology, personality, skill) can do so. There is clear evidence that moral identity does not need to act alone. One person’s fragile network of self-reflections in the table might spin apart as a result of a moral ecology of corruption. Or one can imagine the web of consistencies beginning to come together in a moral ecology that favors particular consistencies, such as restorative justice hearings. Of course, most moral

ecologies contain both integrative and disintegrative influences (e.g. the complexity of a NATO study panel on nuclear technology). Still, one can likely cobble together a modest drive to unity from a wide variety of both internal and external influences. This is likely how most friendships, families, organizations, and societies survive.

4. *A modest drive to unity may be enough to support consistent moral action.*

Moral identity is an important, and even crucial, influence on moral action. But it is rarely the moral hero standing alone in the gap. It provides internal resources for moral action when, because of its relevance in the situation, and sometimes despite the situation, it is called upon by individuals. But it is not always called upon, or it may often be called upon in different forms: the moral identity one has at work may be different from the one most prominent at home. Is consistency among this complexity a vain hope? We think not. There is ample evidence of moral exemplars in a wide range of areas who are thoughtful and mostly consistent in moral action. Moral identity seems to play a central role in their action. We have also reviewed here evidence that variation in moral identity is clearly associated with variations in moral action. So, the mere complexity of the task facing us is not cause for despair. As Frimer and Walker (2008, p. 346) argue, even a modest drive to unity, aided by self-reflection, may be all we need. But the complexity is a caution that we need a better understanding of the task before us. We might instead, reconfigure the metaphor of “bridging the gap” into one of maintaining an ever-changing network or weave of moral commitments and action. Who is the weaver? In Chapter 9, we argue that it is an active but not omnipotent moral self, navigating the world in response to a range of needs, constraints, and moral goals.

5.5.2 Application

“All I knew was that I wanted grace, and that I needed prayer, and that I was helpless without God, and that I wanted to do everything that people did to keep close to him” (Merton, 1948/1998, p. 329). This quote comes from Merton’s autobiography and is included in the first chapter of this book. It is a very personal description of his reasons for joining a monastic order. At least from this report, the entire self is focused on the search for closeness to God. Nothing else matters. Yet Merton became a central figure in the peace movement in the United States in the 1960s. His overwhelming preference was to be a hermit, yet he was rarely allowed this luxury. Instead, he was told by his Abbot to write, since his writing was of such a high quality and so widely sought after. He was so deeply steeped in Christian monastic tradition and practice that even his turn toward those justice issues that were raging at the time (e.g. race and peace) was suffused with symbolism and language from that tradition. In a famous vision he had in 1958, at the center of the Louisville KY shopping district on the corner of Fourth and Walnut, he saw that everyone in the crowd of shoppers had the spark of God in them. “If only everybody could realize this!” he later wrote of this experience, “But it cannot be explained. There is no way of telling people that they are all walking around shining like the sun” (Merton, 1966, pp. 153–154). This mystical language has nothing of human rights or justice language in it, and yet it marks a turning point in Merton’s writing toward issues of justice, race, and opposition

to war. It is a complicated matter to explain how this strange language of people “walking around shining like the sun” connects to these moral issues and to Merton’s social activism. But it is at least clear that Merton’s moral action has different roots than a moral identity centered on principled justice reasoning. These roots are uniquely grounded in a specific tradition of language, narrative, and worldview where reform of the self, otherworldly desire for God, and social activism are often closely intertwined. There are consistency pressures within this moral ecology, but they are (often) not ones that valorize the self. One can multiply examples like this endlessly, showing that moral action is connected with a view of “a person like me in a situation like this” but that, from one actor to the next, this view is rarely the same in specific motivation and structure. And yet, at an abstract level, we will find things in this chapter that help us understand Merton’s moral action (e.g. Maslow and self-transcendence).

5.5.3 Open Questions

1. *How do people navigate the multiplicities of consistency that the literature suggests?*

We have concluded that the moral self is multidimensional and variable over time and domains, and that there are many dimensions on which individuals can evaluate their own and each other’s behavior. We have also suggested, as in some idiographic personality models, that people will establish complex patterns of moral commitment and moral self-reflection that will vary within these dimensions of the self and the moral ecology. It is these patterns that should have some modest relationship to moral action. But how do people navigate this complex inner and outer landscape? Do some navigate it better than others? What does consistency look like in this complex landscape – a straitjacket for action or something necessarily more flexible? Are some moral ecologies better at supporting appropriate navigation? What skills and dispositions might help in getting better at navigation? And how does this navigation relate to moral action? It is in the end an empirical question whether the moral sense is a fragmented collection with no center or a complex weave designed to fit a complex self and environment.

2. *What are the different roles of the self in taking moral action in different moral ecologies?*

How the moral self is constructed and construed, how it relates to cultural values, and how it is evaluated both by the self and others may all vary depending on the moral ecology. The larger cultural surround, organizational culture, and patterns of friendship may interact with each other in shaping moral action and its evaluation. We are beginning to get models of this interactivity of self and cultural surround, but these models have yet to connect to the models and approaches on moral identity, cognition, or emotion in moral psychology.

3. *To what extent is the moral self active in maintaining webs of commitment and action?*

We have suggested it is time to reconfigure the metaphor of “bridging the gap” into one of maintaining an ever-changing network or web of commitments, self-reflection, and action. This reintroduces the self as an active agent in its own moral behavior, thought, and feeling. And it allows us to ask how this activity acts together with the constraints placed upon us given our biological heritage and personality, our skills and habits, and the constraints and opportunities provided

by our moral ecology. This interaction might have them at one time working together (choosing to follow organizational guidance), working in opposition (choosing to reform or subvert unjust laws), or acting in some complex dialectic (advocating that some praiseworthy voluntary action should be made into a requirement). We sometimes use the metaphor of navigation to think about the complexity here, a metaphor borrowed from Aristotle. When and how do actors monitor the interacting currents and winds of consistency? When and how do they, wittingly or unwittingly, modify their ship or steer against the wind? Are they only sailing alone?

4. *How is moral failure integrated into the moral self in a corrective way?*

We have proposed that there is, indeed, some modest drive to unity in self-reflection. But it is crucial that this modest drive responds to situations of inconsistency and moral failure. Weinstein et al. (2013) helpfully propose three integrative characteristics of self-regulation that might support appropriate reflection on moral failure: (1) conscious access to one's emotions, motivations and values; (2) taking responsibility for one's emotions, decisions, and thoughts; and (3) non-defensiveness in response to challenge. As important as these are, they still leave the moral self isolated in its reflection. In Chapters 3, 6, 8, and 9, we have presented other influences and processes that might be helpful in understanding responses to moral failure and in designing small and large environments that are supportive of self-reflective response to moral failure. But we need systematic research in response to moral failure to bring the process into sharp focus.

5.6 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Blasi (1980). "Bridging moral cognition and moral action: A critical review of the literature." The classic review of the problem of the judgment-action gap and the integration of the self as a helpful solution.
- Cervone (2021). "Five paths to personality coherence." A parallel attempt to construct different aspects of how one evaluates personality coherence, and their interactions with situational context. It suggests the available resources and constraints as individuals search for authenticity.
- Frimer and Walker (2009). "Reconciling the self and morality: an empirical model of moral centrality development." An empirical paper that provides evidence that the relationship between the self's interests and moral concerns ideally transforms from one of mutual competition to one of synergy.
- Weinstein et al. (2013). "The integrative process: New research and future directions." A guide to aspects of self-regulation (awareness, ownership, nondefensiveness) that might help in maintaining integrity in evaluation of the self.
- Three models of the self that are multidimensional:
 - McGregor and Little (1998). "Personal projects, happiness, and meaning: On doing well and being yourself."

- McConnell (2010). “The multiple self-aspects framework: Self-concept representation and its implications.”
- Markus and Kitayama (2010) “Cultures and selves: A cycle of mutual constitution.”

References

- Aktipis, C. A., & Kurzban, R. (2004). *Is Homo Economicus Extinct? Vernon Smith, Daniel Kahneman and the Evolutionary Perspective. Advances in Austrian Economics*, vol. 7 (pp. 135–153). London: Emerlad.
- Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: the “other-praising” emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology*, 4(2), 105–127. doi:10.1080/17439760802650519
- Amiot, C. E., de la Sablonnière, R., Terry, D. J. et al. (2007). Integration of social identities in the self: Toward a cognitive-developmental model. *Personality and Social Psychology Review*, 11(4), 364–388. doi:10.1177/1088868307304091
- Andersen, S. M., & Chen, S. (2002). The relational self: An interpersonal social-cognitive theory. *Psychological Review*, 109, 619–645.
- Aquino, K., McFerran, B., & Laven, M. (2011). Moral identity and the experience of moral elevation in response to acts of uncommon goodness. *Journal of Personality and Social Psychology*, 100(4), 703–718. doi:10.1037/a0022540
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440.
- Aquino, K., Reed, A., Thau, S. et al. (2006). A grotesque and dark beauty: How moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. *Journal of Experimental Social Psychology*, 43, 385–392. doi:10.1016/j.jesp.2006.05.013
- Arnold, M. B. (1960). *Emotion and Personality*. New York: Columbia University Press.
- Augustine, M. J. (1986). Zen and Benedictine monks as mythopoeic models of nonegocentered worldviews and lifestyles. *Buddhist-Christian Studies*, 6, 23–49.
- Badhwar, N. K. (1993). Altruism versus self-interest: sometimes a false dichotomy. *Social Philosophy and Policy: Special Issue on Altruism*, 10(1), 90–117.
- Bandura, A. (1986). *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87–99.
- Beall, J., & Restall, G. (2019). Logical consequence. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/logical-consequence/>
- Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, 88(1), 1–45.
- Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Morality, Moral Behavior, and Moral Development* (pp. 129–139). New York: Wiley.
- Blasi, A. (1999). Emotions and moral motivation. *Journal for the Theory of Social Behaviour*, 29(1), 1–19.
- Blasi, A. (2005). Moral character: A psychological approach. In D. Lapsley & F. C. Power (Eds.), *Character Psychology and Character Education* (pp. 67–100). Notre Dame: University of Notre Dame Press.
- Blasi, A. (2009). The moral functioning of mature adults and the possibility of fair moral reasoning. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 396–440). Cambridge: Cambridge University Press.

- Bluck, S., Alea, N., Habermas, T. et al. (2005). A tale of three functions: The self-reported uses of autobiographical memory. *Social Cognition*, 23(1), 91–117.
- Bowman, N. A., Kitayama, S., & Nisbett, R. E. (2009). Social class differences in self, attribution and attention: Socially expansive individualism of middle-class Americans. *Personality and Social Psychology Bulletin*, 35(7), 880–893.
- Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology*, 28(5), 759–775.
- Cervone, D. (2004). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, 56, 423–452.
- Cervone, D. (2021). Five paths to personality coherence: Integrative implications of the Knowledge-and-Appraisal Personality Architecture. *European Journal of Personality*, 36(3), 319–346. doi:10.1177/08902070211015599
- Cervone, D., & Tripathi, R. (2009). The moral functioning of the person as a whole: On moral psychology and personality science. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 30–51). New York: Cambridge University Press.
- Cojuharencu, I., Cornelissen, G., & Karelaia, N. (2016). Yes, I can: Feeling connected to others increases perceived effectiveness and socially responsible behavior. *Journal of Environmental Psychology*, 48, 75–86. doi:10.1016/j.jenvp.2016.09.002
- Colby, A., & Damon, W. (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. New York: Free Press.
- Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, 38, 907–919.
- Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology*, 78(4), 791–808. doi:10.1037/0022-3514.78.4.791
- Davenport, J. (2013). Selfhood and “Spirit”. In J. Lippett & G. Pattison (Eds.), *The Oxford Handbook of Kierkegaard* (pp. 230–251). Oxford: Oxford University Press.
- Day, L. C., & Impett, E. A. (2018). How interdependent self-construal shapes willingness to sacrifice and satisfaction with sacrifice in romantic relationship. *Journal of Social and Personal Relationships*, 1–21. doi:10.1177/0265407517694965
- De Silveira, C., & Habermas, T. (2011). Narrative means to manage responsibility in life narratives across adolescence. *The Journal of Genetic Psychology*, 172(1), 1–20.
- DeWall, C. N. (2009). The pain of exclusion: Using insights from neuroscience to understand emotional and behavioral consequences to social exclusion. In M. J. Harris (Ed.), *Bullying, Rejection, and Peer Victimization: A Social Cognitive Neuroscience Perspective* (pp. 201–224). New York: Springer.
- Ellsworth, P. C. (2013). Appraisal theory: Old and new questions. *Emotion Review*, 5(2), 125–131. doi:10.1177/1754073912463617
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Flanagan, O., Sarkissian, H., & Wong, D. (2008). Naturalizing ethics. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Evolution of Morality: Adaptations and Innateness* (Vol. 1, pp. 1–26). Cambridge, MA: MIT Press.
- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1), 5–20.
- Frimer, J. A., & Walker, L. J. (2008). Towards a new paradigm of moral personhood. *Journal of Moral Education. Special Issue: Towards an Integrated Model of Moral Functioning*, 37(3), 333–356. doi:10.1080/03057240802227494

- Frimer, J. A., & Walker, L. J. (2009). Reconciling the self and morality: an empirical model of moral centrality development. *Developmental Psychology*, 45(6), 1669–1681. doi:10.1037/a0017418
- Frimer, J. A., Walker, L. J., Dunlop, W. L. et al. (2011). The integration of agency and communion in moral personality: Evidence of enlightened self-interest. *Journal of Personality and Social Psychology*, 101(1), 149–163.
- Garff, J. (2013). Formation and the critique of culture. In J. Lippett & G. Pattison (Eds.), *The Oxford Handbook of Kierkegaard* (pp. 252–272). Oxford: Oxford University Press.
- Gelfand, M. (2021). Cultural evolutionary mismatches in response to collective threat. *Current Directions in Psychological Science*, 30(5), 401–409. doi:10.1177/09637214211025032
- Gergen, K. (2000). *The Saturated Self: Dilemmas of Identity in Contemporary Life*. New York: Basic Books.
- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin*, 136(3), 351–374. doi:10.1037/a0018807
- Grube, J. A., & Piliavin, J. A. (2000). Role identity, organizational experiences, and volunteer performance. *Personality and Social Psychology Bulletin*, 26(9), 1109–1119.
- Habermas, T. (2011). Identität und Lebensgeschichte heute. Die form autobiographischen Erzählens. *Psyche*, 65(7), 646–667.
- Haidt, J. (2000). The positive emotion of elevation. *Prevention and Treatment*, 3, 1–5.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, Scherer, K. R., & Goldsmith, H. H. (Eds.), *Handbook of Affective Sciences* (pp. 852–870). Oxford: Oxford University Press.
- Haidt, J., & Kesebir, S. (2010). Morality. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (5th ed., pp. 787–832). Hoboken, NJ: Wiley.
- Hardy, S. A., Bhattacharjee, A., Reed, A. et al., (2010). Moral identity and psychological distance: The case of adolescent parental socialization. *Journal of Adolescence*, 33(1), 111–123. doi:10.1016/j.adolescence.2009.04.008
- Hart, D. (2005a). Adding Identity to the Moral Domain. *Human Development*, 48, 257–261.
- Hart, D. (2005b). The development of moral identity. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (pp. 166–196). Lincoln, NE: University of Nebraska Press.
- Hart, D., Atkins, R., & Southerland, N. (2006). Community service and moral development. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 633–656). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hart, D., & Carlo, G. (2005). Moral Development in Adolescence. *Journal of Research on Adolescence. Special Issue: Moral Development*, 15(3), 223–233. doi:10.1111/j.1532-7795.2005.00094.x
- Hart, D., & Matsuba, M. K. (2007). The development of pride and moral life. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The Self-Conscious Emotions: Theory and Research* (pp. 114–133). New York: Guilford Press.
- Hart, D., & Matsuba, M. K. (2009). Urban neighborhoods as contexts for moral identity development. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 214–231). New York: Cambridge University Press.
- Hartshorne, H., & May, M. A. (1928). *Studies in the Nature of Character: Studies in Deceit* (Vol. 1). New York: Macmillan.
- Heatherington, T. F. (2011). Neuroscience of self and self-regulation. *Annual Review of Psychology*, 62(1), 363–390. doi:10.1146/annurev.psych.121208.131616
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

- Huff, C., & Barnard, L. K. (2009). Good computing: moral exemplars in the computing profession. *IEEE Technology and Society Magazine* (Fall), 47–54.
- Huff, C., Barnard, L. K., & Frey, W. (2008a). Good computing: a pedagogically focused model of virtue in the practice of computing (part 1). *Journal of Information, Communication and Ethics in Society*, 6(3), 246–278.
- Huff, C., Barnard, L. K., & Frey, W. (2008b). Good computing: A pedagogically focused model of virtue in the practice of computing (part 2). *Journal of Information, Communication & Ethics in Society*, 6, 284–316.
- Jennings, P. L., Mitchell, M. S., & Hannah, S. T. (2015). The moral self: A review and integration of the literature. *Journal of Organizational Behavior*, 36(S1), S104–S168. doi:10.1002/job.1919
- Kant, I. (1785/2011). *Groundwork of the Metaphysics of Morals: A German-English Edition* (M. Gregor & J. Timmermann, Trans.). Cambridge: Cambridge University Press.
- Killen, M., & Dahl, A. (2021). Moral reasoning enables developmental and social change. *Perspectives on Psychological Science*, 16(6), 1209–1225. doi:10.1177/1745691620964076
- Kohlberg, L. (1958). The development of modes of moral thinking and choice in the years 10 to 16. PhD dissertation. University of Chicago, Chicago.
- Kohlberg, L. (1963). The development of children's orientations toward a moral order: I. Sequence in the development of moral thought. *Vita Humana*, 6(1–2), 11–33.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive Development and Epistemology* (pp. 151–235). New York: Academic Press.
- Kohlberg, L., Levine, C., & Hewer, A. (1983). *Moral Stages: A Current Formulation and a Response to Critics*. Basel: S. Karger.
- Koltko-Rivera, M. E. (2006). Rediscovering the later version of Maslow's hierarchy of needs: Self-transcendence and opportunities for theory, research, and unification. *Review of General Psychology*, 10(4), 302–317. doi:10.1037/1089-2680.10.4.302
- Koole, S. L. (2009). The psychology of emotion regulation: An integrative review. *Cognition & Emotion*, 23(1), 4–41. doi:10.1080/02699930802619031
- Koole, S. L., Greenberg, J., & Pyszczynski, T. (2006). Introducing science to the psychology of the soul: Experimental existential psychology. *Current Directions in Psychological Science*, 15(5), 212–216. doi:10.1111/j.1467-8721.2006.00438.x
- Koole, S. L., & Veenstra, L. (2015). Does emotion regulation occur only inside people's heads? Toward a situated cognition analysis of emotion-regulatory dynamics. *Psychological Inquiry*, 26(1), 61–68. doi:10.1080/1047840X.2015.964657
- Kuhl, J., & Koole, S. L. (2004). Workings of the will: A functional approach. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of Experimental Existential Psychology* (pp. 411–430). New York: Guilford Press.
- Lapsley, D. K. (2008). Moral self-identity as the aim of education. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of Moral and Character Education* (pp. 30–52). New York: Routledge.
- Lapsley, D. K., & Hill, P. L. (2008). On dual processing and heuristic approaches to moral cognition. *Journal of Moral Education*, 37(3), 313–332. doi:10.1080/03057240802227486
- Lapsley, D. K., & Narvaez, D. (2004). A social-cognitive approach to the moral personality. In D. K. Lapsley & D. Narvaez (Eds.), *Moral Development, Self, and Identity* (pp. 189–212). Mahwah, NJ: Lawrence Erlbaum.
- Lapsley, D. K., & Narvaez, D. (2005). Moral psychology at the crossroads. In D. K. Lapsley & F. C. Power (Eds.), *Character Psychology and Character Education* (pp. 18–35). South Bend, IN: University of Notre Dame Press.

- Lazarus, R. S. (1991a). Cognition and motivation in emotion. *American Psychologist*, 46(4), 352–367.
- Lazarus, R. S. (1991b). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819–834.
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58, 317–344. doi:10.1146/annurev.psych.58.110405.085658
- Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, 5(4), 420–430. doi:10.1177/1745691610375557
- McAdams, D. P. (2006). The problem of narrative coherence. *Journal of Constructivist Psychology*, 19(2), 109–125. doi:10.1080/10720530500508720
- McAdams, D. P. (2009). The moral personality. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology*. (pp. 11–29). New York: Cambridge University Press.
- McAdams, D. P., & Pals, J. L. (2006). A new big five: Fundamental principles for an integrative science of personality. *American Psychologist*, 61(3), 204–217.
- McConnell, A. R. (2010). The multiple self-aspects framework: Self-concept representation and its implications. *Personality and Social Psychology Review*, 15(1), 3–27. doi:10.1177/1088868310371101
- McGregor, I., & Little, B. R. (1998). Personal projects, happiness, and meaning: On doing well and being yourself. *Journal of Personality and Social Psychology*, 74(2), 494–512.
- Merton, T. (1948/1998). *The Seven Story Mountain*. New York: Houghton Mifflin Harcourt.
- Merton, T. (1966). *Conjectures of a Guilty Bystander*. New York: Crown
- Merton, T. (2009). *The Rule of Saint Benedict: Initiation into the Monastic Tradition 4*. Collegeville, MN: Liturgical Press.
- Miller, J. G., Bersoff, D. M., & Harwood, R., L. (1990). Perceptions of social responsibilities in India and in the United States: Moral imperatives of personal decisions? *Journal of Personality and Social Psychology*, 58(1), 33–47.
- Miller, J. G., Goyal, N., & Wice, M. (2017). A cultural psychology of agency: morality, motivation, and reciprocity. *Perspectives on Psychological Science*, 12(5), 867–875.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.
- Moghaddam, F. M. (2005). The staircase to terrorism: a psychological exploration. *American Psychologist*, 60(2), 161–169. doi:10.1037/0003-066X.60.2.161
- Moghaddam, F. M., Warren, Z., & Love, K. (2014). Religion and the staircase to terrorism. In R. F. Paloutzian & C. L. Park (Eds.), *Handbook of the Psychology of Religion and Spirituality* (pp. 632–647). New York: Guilford.
- Moll, J., Krueger, F., Zahn, R. et al., (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences*, 103(42), 15623–15628.
- Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, 81, 33–43.
- Narvaez, D. (2010a). The embodied dynamism of moral becoming: Reply to Haidt (2010). *Perspectives on Psychological Science*, 5(2), 185–186.
- Narvaez, D. (2010b). Moral Complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Narvaez, D., Lapsley, D. K., Hagele, S. et al., (2006). Moral chronicity and social information processing: Tests of a social cognitive approach to the moral personality. *Journal of Research in Personality*, 40(6), 966–985. doi:10.1016/j.jrp.2005.10.003

- Niebuhr, R. (1935). *An Interpretation of Christian Ethics*. New York: Harper & Brothers.
- Nisan, M. (1991). The moral balance model: theory and research extending or understanding of moral choice and deviation. In W. M. Kurtines, J. Gewirtz, & L. Lamb (Eds.), *Handbook of Moral Behavior and Development* (Vol. 3). East Sussex: Psychology Press.
- Nucci, L. P. (2000). The promise and limitations of the moral self construct. Paper presented at the Jean Piaget Society: Society for the Study of Knowledge and Development, Montreal, Canada. <http://tiger.uic.edu/~lnucci/MoralEd/articles/nuccipromise.html>
- Nucci, L. P. (2004). Reflections on the moral self construct. In D. K. Lapsley & D. Narvaez (Eds.), *Moral Development, Self and Identity* (pp. 111–132). Mahwah, NJ: Erlbaum.
- Oliner, S. P., & Oliner, P. M. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press
- Opatow, S. (2001). Social Injustice. In D. J. Christie, R. V. Wagner, & D. A. Winter (Eds.), *Peace, Conflict, and Violence: Peace Psychology for the 21st Century*. Englewood Cliffs, NJ: Prentice-Hall.
- Opatow, S. (2005). Moral Exclusion to moral inclusion: Theory for teaching peace. *Theory into Practice*, 44(4), 303–318.
- Oveis, C., Horberg, E. J., & Keltner, D. (2010). Compassion, pride, and social intuitions of self-other similarity. *Journal of Personality and Social Psychology*, 98, 618–630.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148–158.
- Piliavin, J. A., & Callero, P. L. (1991). *Giving Blood: The Development of an Altruistic Identity*. Baltimore, MD: John Hopkins University Press.
- Plaisance, P. (2014). *Virtue in Media: The Moral Psychology of Excellence in News and Public Relations*. London: Routledge.
- Power, F. C. (2004). The moral self in community. In D. K. Lapsley & D. Narvaez (Eds.), *Moral Development, Self and Identity* (pp. 47–64). Mahwah, NJ: Lawrence Erlbaum.
- Pratt, M. W., Hunsberger, B., Pancer, S. M. et al., (2003). A longitudinal analysis of personal values socialization: Correlates of a moral self-ideal in late adolescence. *Social Development*, 12(4), 563–585.
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1), 1–20. doi:10.1017/s0140525x02000018
- Rasmussen, A. S., & Habermas, T. (2011). Factor structure of overall autobiographical memory usage: The directive, self, and social functions revisited. *Memory*, 19(6), 597–605.
- Reed, A., & Aquino, K. (2003). Moral identity and the expanding circle of moral regard towards out-groups. *Journal of Personality and Social Psychology*, 84(6), 1270–1286.
- Reed, A., & Bolton, L. E. (2005). The complexity of Identity. *MIT Sloan Management Review*, 46(3), 18–22.
- Reynolds, S. J., & Ceranic, T. L. (2007). The effects of moral judgment and moral identity on moral behavior: An empirical examination of the moral individual. *Journal of Applied Psychology*, 92(6), 1610.
- Rudd, A. (2007). In defence of narrative. *European Journal of Philosophy*, 17(1), 60–75.
- Rule, J. T., & Bebeau, M. J. (2005). *Dentists Who Care: Inspiring Stories of Professional Commitment*. Chicago: Quintessence Publishing Company.
- Rusbult, C. E., Kubacka, K. E., Kumashiro, M. et al., (2009). “The Part of Me that You Bring Out”: Ideal Similarity and the Michelangelo Phenomenon. *Journal of Personality and Social Psychology*, 96(1), 61–82. doi:10.1037/a0014016
- Sartre, J.-P. (1943/1984). *Being and Nothingness*. New York: Simon & Schuster.
- Schmidt, L. A., & Poole, K. L. (2020). Quiet strengths: Adaptable introversion in the workplace. In L. A. Schmidt & K. L. Poole (Eds.), *Adaptive Shyness: Multiple Perspectives on Behavior and Development* (pp. 182–197). Cham, CH: Springer.

- Schwartz, S. H. (2016). Basic individual values: Sources and consequences. In D. Sander & T. Brisch (Eds.), *Handbook of Value*. Oxford: Oxford University Press.
- Schwartz, S. H., & Bilsky, W. (1987). Toward a universal structure of human values. *Journal of Personality and Social Psychology*, 53(3), 550–562.
- Schwitzgebel, E. (2019). Aiming for moral mediocrity. *Res Philosophica*, 96(3), 347–368.
- Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and consistency from intra- and inter-personal interactions. *Personality and Social Psychology Review*, 6(4), 316–325.
- Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology*, 67(4), 674–687.
- Shulman, D., & Stroumsa, G. G. (Eds.). (2002). *Self & Self-Transformation in the History of Religions*. New York: Oxford University Press.
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, 4(4), 267–281.
- Skitka, L. J., & Mullen, E. (2002). The dark side of moral conviction. *Analyses of Social Issues and Public Policy*, 2(1), 35–41.
- Snyder, M., & Omoto, A. M. (2008). Volunteerism: Social issues perspectives and social policy implications. *Social Issues and Policy Review*, 2(1), 1–36.
- Stellar, J. E., & Keltner, D. (2014). Compassion. In M. M. Tugade, M. N. Shiota, & L. D. Kirby (Eds.), *Handbook of Positive Emotions* (pp. 329–341). New York: Guilford Press.
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171. doi:10.1016/j.cognition.2013.12.005
- Stroud, S. (2019). Weakness of the will. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/weakness-will/>
- Stürmer, S., & Snyder, M. (2010). Helping “us” versus “them”: Towards a group-level theory of helping and altruism within and across group boundaries. In S. Stürmer & M. Snyder (Eds.), *The Psychology of Prosocial Behavior* (pp. 33–58). Malden, MA: Blackwell.
- Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, 28(04), 531–542.
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58, 345–372.
- Tappan, M., & Brown, L. M. (1989). Stories told and lessons learned: Toward a narrative approach to moral development and moral education. *Harvard Educational Review*, 59(2), 182–206. doi:10.17763/haer.59.2.d364up55vx875411
- Vianello, M., Galliani, E. M., & Haidt, J. (2010). Elevation at work: The effects of leaders’ moral excellence. *Journal of Positive Psychology*, 5(5), 390–411. doi:10.1080/17439760.2010.516764
- Vignoles, V. L. (2017). Identity: Personal AND Social. In *Oxford Handbook of Personality and Social Psychology*. Oxford: Oxford University Press.
- Walker, L. J. (2004). Gus in the gap: Bridging the judgment–action gap in moral functioning. In D. Lapsley & D. Narvaez (Eds.), *Moral Development, Self, and Identity* (pp. 1–20). Mahwah, NJ: Lawrence Erlbaum.
- Walker, L. J., & Frimer, J. A. (2009). Moral Personality Exemplified. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 232–255). New York: Cambridge University Press.
- Walker, L. J., & Frimer, J. A. (2015). Developmental trajectories of agency and communion in moral motivation. *Merrill-Palmer Quarterly*, 61(3), 412–439.
- Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review*, 8(3), 281–307. doi:10.1207/s15327957pspr0803_4

- Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2013). The integrative process: New research and future directions. *Current Directions in Psychological Science*, 22(1), 69–74. doi:10.1177/0963721412468001
- Weng, H. Y., Fox, A. S., Shackman, A. J. et al., (2013). Compassion training alters altruism and neural responses to suffering. *Psychological Science*, 24(7), 1171–1180. doi:10.1177/0956797612469537
- Wong, D. B. (2015). Growing virtue: The theory and science of developing compassion from a mencian perspective. In B. Bruya (Ed.), *The Philosophical Challenge from China* (pp. 23–58). Boston, MA: MIT Press.

Skills and Knowledge

A good will is not good because of what it effects or accomplishes, ... but good just by its willing

(Kant, 1785/2011, p. 17)

If we are Kantians, when we say someone “means well” we focus our attention on their intention. But in normal language having “good intentions” is more often an excuse for the *failure* to achieve the stated goal – they *meant* well. This excuse points to an influence on moral action that is rarely referenced. Every moral action case we reviewed in the Introduction required more than simple good intention. Each one required some skill in achieving the goal, even to the level of expert knowledge. Though ethics courses primarily teach moral reasoning, the needed skill is rarely that of moral judgment to determine what is right (Lapsley & Narvaez, 2005). It is more often some mixture of skill in recognizing the problem, framing it, planning for action to resolve it, and implementing those plans (Moulton et al., 2007; Narvaez, 2005).

Ethical education guidelines for middle school (Narvaez, 2006) and college (Callahan, 1980) and guidelines for ethics instruction in computing (Huff & Martin, 1995), dentistry (Bebeau, 1994), psychology (De las Fuentes, Willmuth, & Yarrow, 2005), science (Mumford et al., 2008), and many other areas make explicit links between skill and knowledge and ethical competence.¹ Compared to the proportion of work in moral psychology that investigates moral skill, the conspicuous attention paid to skills in ethics guidelines seems excessive. This is in part because moral psychology has mostly concentrated on only those skills involved in

¹There is some disagreement over whether knowledge is required to have a virtue (Winter, 2011). The discussions here about automatized skill and implicit knowledge may help resolve this controversy.

moral judgment (Lapsley & Narvaez, 2005) to the neglect of skill in other aspects of moral action.²

In this chapter we will first look at both general and domain-specific skills that are associated with moral action and then review two research areas with implications for moral skill: expertise and habit. After surveying the range of human skilled behavior that expertise and habit cover, we conclude that principled, conscious, moral decision-making is probably not the primary path to moral action. And that it is likely the interface between automatic moral action and the shift to conscious reflection when it is required is a critical mark of expertise in any particular moral domain.

6.1 General Skills

The skills associated with moral action can span the entire range from perception to judgment, planning, action implementation, and reflection. Most work in psychology has been done on those skills relevant to moral judgment. We will look at this literature in Chapter 7. Here, we look at other general skills, including self-reflection, self-regulation, and moral attentiveness. These are skills that are assumed to underlie all moral action. Finally, we look at the special skills and knowledge that belong to moral action in specific domains.

We are interested in more than what is covered by the field of self-regulation. But even if we were to limit ourselves to the traditional topics in self-regulation, we face a distinct challenge. After more than a century of investigation of self-regulation, beginning with Freud's (1895/2000) models of internal processes,³ "the self-regulation literature commonly is said to lack theoretical order" (Cervone et al., 2006, p. 333). In a short chapter, we cannot pretend to impose that order. Instead, we present here only a selective review most relevant to moral action.⁴

We have already noted that these skills of moral action need to be understood on various levels. A common mistake is to focus primarily on *conscious* self-control. Such approaches assume a central executive system that consciously guides thought, emotion, and behavior. But there is too much to do in everyday living for us to regulate it all consciously: "filtering, organizing, and integrating the vast array of different feelings, thoughts, needs, motivations, goals, norms, expectations, etc." on an ongoing basis, and also being reflective about those processes (Kuhl & Koole, 2004, p. 415). And even "executive function" is actually a diverse set of functions each of which develop in domain-dependent ways dependent on "knowledge, beliefs, norms,

²And this narrowing of the literature has in turn been supported by the structuralist foundations of Piagetian and Kohlbergian moral psychology (Narvaez, 2005). This old foundation assumes that the conceptual development that supports all morality is separable and unique from other forms of learning, such as expertise. This assumption is no longer widely accepted (Hatano & Inagaki, 2000, p. 267). Rethinking morality as skill will require a rethinking of these Piagetian foundations of moral psychology.

³For modern explanations of what is now called self-regulation, one can arguably go back even further than Freud, at least to Søren Kierkegaard (1849/2004) and his investigation of "inner movements." More ancient distinctions go back to the stoics (Nussbaum, 1996).

⁴See Vohs and Baumeister (2016) for a more comprehensive review of this literature.

values and preferences” (Doebel, 2020). A set of theoretical approaches that take this complexity into account is “self-organizing dynamic systems” (Carver & Scheier, 2002). Instead of assuming a hierarchical, conscious system, these approaches envision multiple psychological systems operating semi-independently and interactively, at both conscious and nonconscious levels, both competing and cooperating.⁵ Integration among these systems is crucial to proper operation of all of the skills (including even the specific domain-relevant skills) we discuss in this chapter (Benita, 2020; Weinstein, Przybylski, & Ryan, 2013). We cover aspects of integration in Chapters 4, 5, and 9. There is a clear commonality here with the two-process systems we cover in Chapter 7, though the number of systems is multiplied. In this chapter, instead of concentrating on cognition, we are primarily looking at how the skills of self-regulation influence and guide our moral action.⁶

6.1.1 Self-Reflection

By moral self-reflection we mean the (mostly) conscious reflection and evaluation of the self and its skills, knowledge, emotions, values, goals, etc. that is done in the service of becoming a different kind of person or of maintaining consistency within one’s self.⁷ Narvaez (2010) calls this aspect moral metacognition, and treats it as essential to the high-level monitoring needed to plan and sustain moral action (see also Kuhl & Koole, 2004; Walker & Frimer, 2009). As the “meta” suggests, these are about higher processes that are more conscious, controlled, and based in reasoning and reflection. This does not rule out an emotional aspect to evaluation, nor does it suggest that some evaluations (e.g. automatic reappraisal of emotions) might be so well practiced that they become automatic.⁸

In the longstanding person–situation debate, a simplistic view is to ask whether the situation or internal processes of the person has more control over behavior.⁹ A more sophisticated approach is to ask when, or under what conditions, the situation or the person is more in control. Impulsive, intuitive action is often more controlled by the situation (Hofmann, Friese, & Strack, 2009). But this can be somewhat regulated by individuals taking precaution to remind themselves of prior commitments (Gollwitzer et al., 2011). Indeed, this is often what happens in the reappraisal of emotion: an initial emotional reaction is tempered by further reflection (Feinberg et al., 2012). For our purpose, one may think of it as reappraisal focused particularly on self-reflective emotional responses (e.g. Am I being proud in the right way?).

Chapter 5 provides us with an idea of the structure and complexity of this process of self-reflection. Table 5.1 in that chapter provides a multidimensional approach to

⁵This approach sometimes seems quite Freudian, and prone to the same difficulties of taking proposed theoretical mechanisms and making them more real than they deserve.

⁶See John Davenport’s excellent (2007) book on the will for a philosophical approach to this area.

⁷We list self-reflection independently of the self-regulation category because its clearly conscious status – at least within the Kuhl and Koole (2004) taxonomy we are using – likely excludes it from lists of self-regulation skills.

⁸See Chapters 7 and 8 for suggestions about how this might happen.

⁹See Chapter 4 for this debate.

how one might evaluate one's actions in the light of various standards. Thus, what one might think is a simple issue of evaluation (Is my action moral?) can in fact become quite psychologically complicated.

Wrzus and Roberts (2017) provide a systematic way to understand how reflection on current experience can be transformed into personality change (Wrzus & Roberts, 2017). The core of the model is the transformation of immediate evaluative reactions to one's behavior (e.g. I was not kind) into *desired* expectancies for behavior (I should be more kind) in similar situations in the future. This transformation is accomplished at both an explicit and an implicit level. At the *explicit* level, individuals consciously reflect on the triggering situation, cognitive and emotional state, and immediate reaction in a particular incident, or they later remember that situation and consciously reflect on it. I might, for instance, remember an action, be proud, and resolve to repeat it. Or when I fail to repeat it under relevant circumstances, I might immediately feel guilt that motivates evaluative thought. In each of these examples, I might at the same time *implicitly* be forming associations of emotions, judgments, and perceptions that can build up over time as I attempt to maintain consistency in my behavior.

This change (or stability) in my evaluative and learning processes then shapes my expectation for the next occurrence of a relevant triggering situation, which in turn shapes my state and behavioral expression of that state. Over time, and many repeated cycles of these evaluative and learning processes, this may result in the shaping (formation) of my behavioral expressions and immediate reactions to them. This long-term cyclical process can thus produce personality/character change.

An important point about this model is that many "exit points" from the evaluative process exist in each daily experience. Each exit point allows for the individual to simply confirm that their expectations and reactions do *not* need to change. To confirm that they do not need to become more moral. Suppose I want to become a more generous person. Even if that is my intention, there are many ways for me to fail to achieve it. In my daily (less generous) life, situations that unequivocally push for a change in my generosity may arise too rarely to help me practice the appropriate reactions and judgments and to reflect on that practice. Even when they do arise, I might not notice that a situation is relevant to *my* generosity. Or for a variety of reasons (available excuses, match to my expectations, etc.) I may not experience any of the moral emotions or judgments that might lead to generosity, or I could notice the situation, react to it, and judge my reaction adequate or at least excusable (thus not leading me to form an expectation that I should be more generous later, in similar situations).

Because of these many exit points, it is easier to maintain current personality characteristics than it is to change them: the accumulation of the exit points works toward maintaining stability.¹⁰ But persistent, systematic changes in the situation or consistent practice of the expectations or evaluations *can* produce personality change. Thus, substantial change in moral formation can occur, even in middle and late adulthood, if (1) enough opportunities occur that challenge existing frameworks; (2) the person

¹⁰This model of many exit points is reminiscent of the Latane and Darley (1970) model of the numerous reasons that a bystander may not help in an emergency situation: not noticing, interpreting it wrongly, etc.

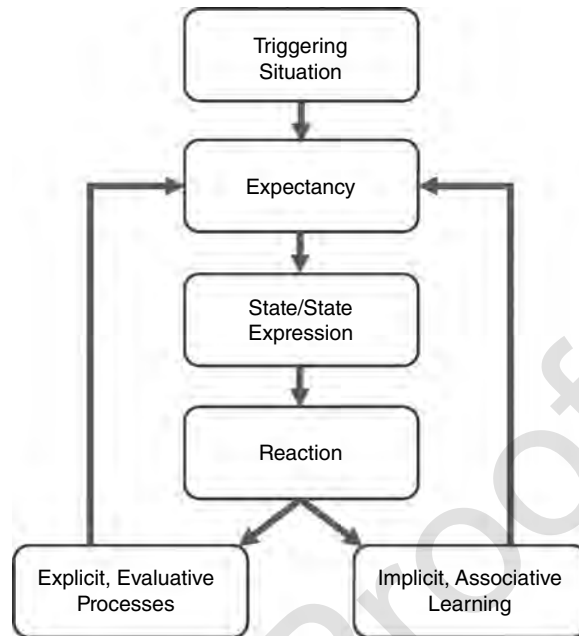


Figure 6.1 A vastly simplified version of the Wrzus and Roberts (2017) model.

experiences the states as novel and calling for self-reflection; and (3) this happens frequently enough to become engrained (Wrzus & Roberts, 2017).¹¹

This framework suggests how, in daily living (see Figure 6.1), consistent practice of shaping actual reactions and expectations in the image of one's desired reactions can shape moral formation.¹² It recognizes that this change occurs at both the explicit, conscious level (e.g. using self-reflection to alter goals, expectations, reactions) and on an implicit, nonconscious level (e.g. in learned associations or trained skills). And it proposes a set of conditions that lead to the most effective and lasting change. The impetus for such change can come from a variety of influences. It can come from consistent personal intention to change, from consistent social pressure, or from life events that exert consistent pressure on the person or some combination of these and other influences. Indeed, the model is so open to different kinds of influences and changes, that it is somewhat unclear exactly what is changing over time. This may be the necessary price we pay for such multidimensional models: what is changing may differ from time to time and person to person. Tracking this complexity will require equally complex idiographic personality models like those we cover in Chapter 4.

¹¹We will suggest in Chapter 9 that some traditional approaches to moral formation try to produce moral ecologies that meet these criteria (see e.g. Chapter 9, Section 9.1.3 on Monastic Formation and Section 9.1.1 on *Bildung*).

¹²Note that underlying this process are many of the skills of self-reflection, self-regulation, emotion regulation, moral sensitivity, etc. that we cover in this chapter.

One can find similar but domain limited psychological models of evaluative moral formation in clinical psychology (Norcross, Krebs, & Prochaska, 2011), religious conversion (Rambo, 1995), and volunteerism (Snyder & Omoto, 2008). These alternative models provide suggestions for the richness and variation of the processes in different domains. The Wrzus and Roberts (2017) model is directed at the most abstract level of the processes and one may find the processes change dramatically in different moral ecologies (e.g. in more wicked or kind environments, see Section 6.3.1).

6.1.2 Self-Regulation

The Mothers of the Plaza de Mayo were focused on calling attention to the plight of their missing children in Argentina.¹³ Even in the face of severe repression (several founders were themselves disappeared), they brought their cause to international attention and played a significant part in the eventual change of government. This is a story of extraordinary courage but also extraordinary skill (Thalhammer et al., 2007). None of this could have happened if the founders and their followers had not been able to organize themselves, plan and execute complex schemes, and adapt to disappointment, fear, and changing conditions. In addition to skill associated with community organizing, many of these skills are more basic strategies of self-regulation.

A comprehensive review of self-regulation could fill a book (and has, Boekarts, Pintrich, & Zeidner, 2005). Suffice it to say that ability in each of these aspects is predictive (or perhaps constitutive) of mature functioning in life (Frimer & Walker, 2008; Mischel, Shoda, & Rodriguez, 1989) and can be increased through practice (Baumeister, Gailliot, & Tice, 2009; Förster & Jostmann, 2012; Lord et al. 2009). We use the term self-regulation to gather together a set of explicit and implicit skills in guiding the self in daily tasks.¹⁴ We will focus on three different aspects of this set of skills:

- The *self-control system*, that set of skills and associated standards, personality traits, etc. that help the person to focus on the task at hand, and to avoid distraction and temptation (Kuhl & Koole, 2004).

¹³See the Introduction for an overview of this movement.

¹⁴There are numerous other models of self-regulation. Some are, like Kuhl and Koole (2004), more friendly to the idea of some complicated version of “free will” while others are more mechanistic. But almost all reside within the domain of what philosophers call *compatibilism*, the idea that free will can be compatible with deterministic causality in human action. That is, most recognize deterministic limits on what can be “free” about free will, but also recruit deterministic processes (e.g. habit formation) in interaction with planned action to explain how free will is implemented in human (and perhaps other animal) action. See Kuhl, Quirin, and Koole (2021) for a recent overview in humans. See de Waal (2009) and Chapter 1 for the argument that extends these processes to some other animals.

- The *self-maintenance system*, that set of skills and associated standards that help direct the person in daily life to take action toward goals, standards, and values that matter to them (Kuhl & Koole, 2004).¹⁵
- *Emotional self-regulation*, a set of skills that overlaps significantly with each of the first two systems and helps to motivate and structure the interaction between them and other aspects of moral life, such as self-reflection and self-formation.

6.1.2.1 Self-Control The self-control system has the function of inhibiting impulsive action and focusing on currently active goals. It is often reduced to what is misleadingly called “willpower.” The “willpower” metaphor suggests what has been called the *ego-depletion hypothesis*: that there is a limited reserve of such power, that it operates the same in all domains, that it always involves conscious focus and effort, that some people have more of it, and that practice can help improve it (Baumeister et al., 2009; Baumeister & Tierney, 2011; Weir, 2012). The effect is in dispute (Frieze et al., 2019; Vohs et al., 2021) but may survive in some altered form (Lin et al., 2020). Still, ego-depletion (and willpower) is not all there is to self-control. Calling it the self-control system opens consideration for other processes that operate differently but serve the same function.

For instance, in a meta-analysis of 120 studies on self-control, de Ridder et al. (2012) found dramatic differences in the effectiveness of self-control in different domains. Self-control’s effects on eating and weight were quite small, its influence on work and school were strong, and its effects on interpersonal goals (e.g. relationship commitment, parental supportiveness) were in between.¹⁶ They also found that people who are better at self-control are not, by and large, better at resisting temptation (the classic willpower task) but more skilled at forming and breaking habits. For these people, self-control is operating at the interface between system 1 and system 2 activity, between habitual, automatic action and conscious guided action.¹⁷ Thus, successful self-control is strongly linked to using this skill to *avoid* temptation, rather than to *resist* it (Ent, Baumeister, & Tice, 2015), and this may better be accomplished at the implicit, nonconscious, habitual level, guided by conscious efforts to cultivate (and weed)

¹⁵For organizational purposes, we use the Kuhl and Koole (2004) distinction between a system that is primarily about avoidance of distraction and temptation (self-control) vs. a system that is primarily about achieving goals. We do not follow their theoretical approach in all its complexity but do adopt their suggestion that affect regulation is central in moderating among the various systems. One reason to use their approach is its openness to *nonconscious* ways of regulating emotion. We also emphasize various *conscious* meta-reflective aspects in this picture. These distinctions are shared among many theorists and are at least useful in bringing some order to the theoretical thickets of work in this area. There are many alternatives to the Kuhl and Koole (2004) approach. One advantage in adopting it in a book that uses virtue as a model is that it takes the notion of will seriously and does so in a way that allows for nonconscious volitional action. For alternative approaches, see Cervone et al. (2006); Shoda, Lee-Tiernan, and Mischel (2002); Ryan and Deci (2003); or Hofmann et al. (2009).

¹⁶It is unclear where morality would fall in this set of domains.

¹⁷Remember that system 1 is the one in which habit, implicit judgment, and nonconscious control of action are occurring, while system 2 involves conscious, effortful, reasoning, planning and implementation. See Chapter 7 for a review of these systems.

one's profile of habits.¹⁸ This is the origin of the standard advice that consistent weight loss is best effected by lifestyle change (the change of habits, surroundings, shopping, etc.) rather than by simple willpower (only eating very small portions of the cake and pastries that fill the refrigerator) (Rothman, Sheeran, & Wood, 2009).

Thus, at its best, self-control functions to avoid distraction and inhibit impulsive action. But it is most successful when it operates to *avoid* temptation (rather than to resist it); when it avoids the need to exercise willpower. One establishes eating habits that avoid certain foods, habits of friendship that avoid opportunities to practice racial prejudice, habits of association that encourage compassion, or habits of giving that encourage generosity. This strategic habit formation is not easy to do, and it is hedged around with the difficulties of situational pressure, opportunity, and available skill. However, deployed with the appropriate guidance of self-reflection, it helps the individual to avoid distraction and to focus on desired activities and outcomes. It merges the task-focus of self-control with the goal-focus of self-maintenance.

6.1.2.2 Self-Maintenance The self-maintenance aspects of self-regulation help focus the person on activities that are intrinsically rewarding or “congruent with a multitude of the person’s inner values and autobiographical experiences” (Kuhl & Koole, 2004, p. 416). It is less concerned with *not doing* (self-control’s domain) and more concerned with acting in the service of chosen goals, values, narratives, etc. It is positive doing. In our example of the Mothers of the Plaza de Mayo, it is less about resisting fear and more about engaging the community in action, or pursuing the goals of the movement. The difficulty, of course, is that there is indeed a multitude of such goals, including social disagreement about them. And many of these disagreements are deeply emotional and central to group identity and intergroup relations (Goldenberg et al., 2020; Porat et al., 2020; E. R. Smith & Mackie, 2008). The self-maintenance aspect of self-regulation must negotiate its way among these goals and around the numerous obstacles in the way (Duckworth et al., 2007). It is the process of everyday negotiation of action in the service of multiple goals that constitutes self-maintenance.

The self-maintenance system operates by integrating multiple aspects from cognitive and affective subsystems to support the intuitive control of goal-directed behavior, control that is purposeful but usually not conscious. It is partly conscious in making short-term plans, but the cognitive structures that need to be accessed even in a normal conversation with friends (ideas, goals, values, expectations, norms, situational knowledge, etc.) are too extensive to be completely contained within conscious experience. If one were to constantly monitor these during conversation, one would become lost in thought. Conversation or other daily tasks only work because these networks are subconsciously explored and, if possible, integrated when there are conflicts. Thus, the self-maintenance system helps to guide, shape, and maintain the engagement of the person in his or her daily life tasks.

So, we have self-control aspects of self-regulation that focus on active suppression of action (e.g. not interrupting in the conversation) and self-maintenance aspects that

¹⁸Long practice of particular actions in religious contexts may make the burden of moral commitment lighter than it might otherwise be (Koole et al., 2010)

guide action (e.g. listening carefully to another person).¹⁹ Then there are conscious self-reflective aspects that focus on how one is meeting goals and standards. Do all these aspects run in parallel or do we switch among them? We really do not have enough data or theory to know. In some domains, like conversation, there may be some parallelism, but in others, such as skilled musical or surgical performance, the systems may be distinct. But one thing that does seem to be true is that emotion (and thus the emotion-regulation system) is involved in signaling and negotiating the need to switch among systems and strategies (Alter et al., 2007).

6.1.2.3 Emotion Regulation When, while performing some action, it seems things are not going right, not on track to meet the goal, one often has a negative emotional experience. This emotion serves as a signal that something is not working in one's action plans and performance (Alter et al., 2007). It signals the need to search for options about why the uncomfortable feeling of "it is not working" has surfaced. It acts as an internal signal that something needs adjusting in order to obtain goals. Thus, our emotional life and its regulation are central in helping us to become better at attaining our goals, and among them, our moral goals.

But emotion regulation is a broader set of processes than simply those that serve us in pursuing our goals. We regulate our emotions whenever we find that we want to redirect the spontaneous flow of our emotions.²⁰ Emotion regulation in the moral domain, then, at the least involves attempting to increase or decrease any of the moral emotions we list in Chapter 8. We might want to increase feelings of compassion to help us become a better person; or decrease feelings of compassion because they make us so uncomfortable. We may want to increase feelings of anger to make an impression in a protest against injustice; or decrease anger to facilitate our interaction with coworkers or family. And we do this for various reasons and using a range of strategies. Thus, there will be a range of motivations and strategies for regulating moral emotion. In an attempt to systematize the complexity here, Koole (2009) proposes three different functions that emotion regulation can serve: goal achievement, hedonic need fulfillment, and personal balance maintenance. Focused, appropriate,

¹⁹These two examples provide a nice picture of how the two "systems" can be distinguished. The phenomenological experience of *trying not to interrupt* is one of refraining from acting, even though one wants to. The intentionality is one of not doing something. The emotional concomitants may include frustration, and even a longing concentration on the forbidden desire. The experience of *listening intently* may also inhibit interruption, or at least transform it into backchannel expressions like "yes, uh-huh" that focus both speaker and listener on the conversation. The phenomenological experience is also different. Not interrupting is consciously frustrating, while listening intently may be one of engagement or even flow.

²⁰One should be suspicious of thinking that emotions are spontaneous things that simply happen to us. Deeply ingrained patterns of emotional responding can feel spontaneous and still be both learned and changeable with practice. But talk about "regulating" assumes there is some process in which we want to intervene. These prior processes that produce or anticipate emotion are often called emotional sensitivity (Koole, 2009) or in the moral domain, moral imagination, ethical sensitivity, or habituated empathic concern (Narvaez, 2010). Thus, those "spontaneous" emotions can themselves be trained, as can the regulation of the spontaneous emotions as they arise.

and balanced emotions are central to each of these functions. Regulation is achieved for each function by means of three different systems – guided attention, knowledge (reasoning and reappraisal), and body systems (manipulating food intake, exercise, attention to emotion expression) – that are engaged in that regulation. This produces a relatively complex grid of three functions crossed with the three systems that serve each function. And as always, the actual processes may not allow for such complex and clear distinctions to be made: systems and functions overlap.²¹

Meditation and mindfulness training can target the attention, knowledge, or body systems depending upon the form (Garland et al., 2015; Hölzel et al., 2011; Kok & Singer, 2017). There is some evidence that mindfulness practice leads to increases in compassion (Hölzel et al., 2011) or what Narvaez (2010) calls “habituated empathy” with changes, again, in all three systems: one becomes more aware, thoughtful, and responsive in those emotions associated with compassion. Mindfulness training also brings awareness of subjective experience (again in all three systems) that can drive behavior change by producing a change in how one perceives, values, and reacts to one’s habits, enhancing the likelihood of desirable change (Ludwig, Brown, & Brewer, 2020). Van Gordon, Saphiang, and Shonin (2021) provide a review of the successes and controversies of meditation including concern over “McMindfulness,” and possible adverse effects. Some of these difficulties come from separating religiously and socially embedded practices from the practices and stances that originally surrounded and supported them (Condon & Makransky, 2020).

Though this presentation of types and systems seems to provide clear guidelines for separating concepts, it is, as always, difficult to make the distinctions in the everyday world of moral action. Several of the strategies target multiple systems, and the types show conceptual and practical overlap. The categories also miss the social nature of emotion-regulation: People often proactively control emotion by creating, modifying, or avoiding particular situations, thus intervening in the process before it has a chance to begin. Or they spend time talking with others, regulating their emotion by sharing it in conversation (Wagner et al., 2015).

Two final points about the skill of emotional regulation: First, changes toward more flexible and appropriate emotional regulation are clearly possible in adulthood. It has been shown that clinical psychological approaches like emotion-focused therapy (EFT) can produce better emotional responding from clients with a range of emotional difficulties (Angus et al., 2015). Mindfulness and meditation programs show improvements in emotional functioning (Kok & Singer, 2017), gratitude training (Bono, Emmons, & McCullough, 2012), and various political consciousness-raising groups (Keane, 2015) all show the processes through which adults can change the ways they emotionally respond to moral issues. Second, as many of the examples show, this kind of emotional development does not simply mean restricting one’s emotional range. It can also mean deepening it so that one responds with more complexity and skill to situations (Frijda & Sundararajan, 2007).

²¹See Koole (2009, p. 11 ff.) for an overview of these difficulties and detailed examples and research on the functions and systems.

Self-regulation is a complex and ever-changing dance among self-control, self-maintenance, and various strategies of emotional regulation. All these aspects of self-regulation, in all their complexity, are skills that contribute to success in a variety of domains, and thus they can be practiced and improved over time with guided practice. We simply do not know enough yet to say which sets of skills are more easily improved, or how this relates to the moral domain in which they might be exercised.

6.1.3 Moral Attentiveness

There is a skill involved in being regularly aware of and ready to respond to morally relevant situations. Narvaez (2010) calls this cognitive chronicity, taking the concept from the social cognition literature about mental models that are complex, detailed, and *chronically* accessible (Higgins, King, & Mavin, 1982). This is a skill aspect of what, in Chapter 5, we call moral identity. Two of the defining characteristics of having a moral identity are that: one has a detailed and complex knowledge of the domain;²² and this knowledge affects the sensitivity of the individual to relevant issues in their environment.²³ Acquiring such skills is obviously part of the task of growing up (Eisenberg, 2005) and some adults clearly excel in this (e.g. Colby & Damon, 1992). In work on leadership in the military, Hanna and Avolio (2010) have called this cluster of skills and knowledge *moral potency* and shown its relationship to ethical values and behavior among soldiers. In work on organizational teamwork, Reynolds (2008) has used a similar concept, *moral attentiveness*, and shown its relation to recall of moral action, moral awareness, and moral behavior in teamwork as rated by teammates. The research on this skill is still somewhat scattered, but the cognitive chronicity approach of Narvaez (2010) provides a useful way of identifying its components and developmental trajectory.

6.1.4 Moral Imagination

Moral imagination is the skill needed to imagine alternatives after one has noticed that there is some moral choice to be made. It is a term used more in business ethics (Preti, 2014; I. H. Smith et al., 2023) and technology ethics (Crowell, Narvaez, & Gomberg, 2004; Friedman & Hendry, 2019; Pritchard, 2001) than in psychology, where the idea is fragmented into theory of mind (Gray, Young, & Waytz, 2012; Hudac & Sommerville, 2020; Wellman, 2018), empathy (Decety, 2021; Depow, Frances, & Inzlicht, 2021; Fowler, Law, & Gaesser, 2021), and option choice and evaluation in decision-making (Galotti, 2007; Morris et al., 2021). It is limited by goal and motivation, as we characterize it in Chapter 9, Section 9.3.1.3 (Opatow, 2010). Part of the point of moral exclusion is the removal of the need to exercise moral imagination on their behalf. Moral imagination is also quite bounded

²² But remember that domain can be quite limited. Morality (and moral identity) may be limited to only one aspect of an individual's life, and this can result in inconsistencies between domains, such as work, charity, and family (Colby & Damon, 1992).

²³ The third aspect of moral identity is that it is closely linked to the individual's view of the self (Lapsley, 2008).

by knowledge and expertise in the specific culture, domain, or situation as we discuss in the next section.

6.2 Domain-Specific Skills

The skills covered so far seem relevant in a wide variety of situations and domains. But there is ample evidence that these skills alone are not sufficient when one is participating in a particular domain of life. One needs skills relevant to each domain. It is possible that the general moral skills may be prerequisites for, or lead individuals to more quickly acquire, the domain-relevant skills.²⁴

Obviously, domain-relevant skills will be particular to each domain, and since the domains are legion in number, so too would be the skills. Generalizing across domains, one might find similarities in these skills, but these similarities would need to occur at a higher level of abstraction, and the skill may not be able to be taught at this level of abstraction. This is interesting speculation, but there is very little work even on domain-relevant skills to support it. There are some particular examples of what we might mean by domain-relevant skills. For instance, when the philosopher Michael Pritchard interviewed moral exemplars in the engineering professions, he found that the virtues they listed contained some standard suspects (e.g. honesty, courage) but also quite specific skills, such as “a habit of documenting work” (Pritchard, 1998, 2001). Moral exemplars in computing make little distinction among moral, social, and technical skills, often claiming that all were interrelated (Huff & Barnard, 2009). And experts in the ethical issues of computing use domain related knowledge (e.g. data privacy) far more than novices do (Keefer & Ashley, 2001). One presumes that expertise consists in part of having these intermediate knowledge structures available and ready to use. Work in addiction recovery (Bernheim, 2004) shows that specific skills of understanding and managing susceptibility to relapse are crucial in recovery. Skills like social networking and negotiating are what Thalhammer et al. (2007) refer to as internal resources in their model of what influences the action of courageous resisters, such as those who sheltered victims of the holocaust. Snyder and colleagues’ model of volunteering suggests that one of the reasons people volunteer is to learn skills they might not find elsewhere (Snyder & Clary, 2000; Snyder & Omoto, 2008) and have a chance to practice them to increase expertise.

Thus, there is something special about the domain-specific knowledge that experts bring to a (moral) task that matters. In Section 6.3, we try to uncover what that is.

²⁴This is similar to Aristotle’s thought when he says that one must have had the right kind of upbringing before one can learn the virtues (Aristotle, 1941, NE 10.9 1179b4–31). But we do not know of any research that would document this.

6.3 Expert Performance

We can think of excellence in moral acting as the achievement of expertise in a domain (Dreyfus & Dreyfus, 2004; Narvaez, 2005, 2010). The model of moral action as expertise conforms well to the recently resurgent idea of moral action consisting of the exercising of the virtues (Annas, 2006; Anscombe, 1958; MacIntyre, 1981). Like virtue, expertise is a characteristic of the individual and can be held at various levels of excellence. Like virtue, expertise consists of habitual and even automatic ways of performing certain tasks, and these ways can be acquired through practice. It is likely that descriptions of an individual as “having” a virtue (e.g. honesty) are actually descriptions of a characteristic way of acting well (acting in an expert manner in regard to being honest), rather than descriptions of a characteristic the individual “has” that causes virtuous action.²⁵

The expertise model of morality thus sees consistent moral action as the use of skills to strive for excellence with regard to particular (moral) goals in a particular domain. The caveat regarding the nature of goals is important here. It has been shown, for instance, that mentoring in service of exclusively success-oriented goals can lead young scientists to actually care less about moral issues, and be less moral in their practice of science (Anderson et al., 2007). As discussed in Chapter 5, one can have many other goals than moral ones, including enjoyment, sociality, and achievement (McGregor & Little, 1998). All people, even psychopaths (Skeem et al., 2011), have some of the skills associated with achieving moral goals, and need them in order to navigate our shared social life. But moral exemplars, individuals recognized for excellence in striving after moral goals, have achieved high levels of expertise in these skills.

Narvaez (2010) makes the analogy with music. Everyone appreciates and participates in music to some extent, but few of us have extensively practiced both the appreciation and performance of music: these are experts. The literature on expertise presents some consensus on the differences between novices and experts who:

1. Have more and better organized domain knowledge (declarative, procedural and conditional).
2. Perceive and react to the world differently, noticing detail, risk, and opportunity.
3. Act in the domain in an automatic and effortless way, when things are normal.

Common to all the models is the idea that expertise can be learned over time, with practice and appropriate feedback (Dane, 2010). This learning arc was noticed long ago by Zeami (Motokio, 1400/1984) who cataloged the developmental progression of expertise in his medieval treatises on Japanese *Noh* theater. According to Zeami,

²⁵Hulsey and Hampson (2014) have proposed a model of virtue as composed of *habits and expert skills*. They combine these under the idea of *Habitus*, borrowing from both Thomas Aquinas and the early emotion researcher Magda Arnold (Arnold, 1960) who based much of her theorizing on Aquinas’ *Habitus*. Their model is consistent with much of what is presented in this text.

the rules that the beginner obeys (*Shu*) are “beautiful fictions,” and with practice the novice can advance to the stage where the practitioner can detach (*Ha*) from the rules and break them intentionally in the right circumstances. True experts transcend (*Ri*) the rules, and even though they may seem to violate them, their action can propose the modification of the rules (Murata, 2010). This *Shu-Ha-Ri* progression has been taken up in the martial arts and in computer science (Boehm & Turner, 2003) to describe levels of mastery.

But can we really speak of expertise in the moral domain? Narvaez (1998) compared moral experts and non-experts and showed the experts behaved like experts in other well-studied domains (e.g. chess). Given complex moral narratives to read, they gave more explanation and total expressions, showed deeper understanding and engagement in texts, disagreed with the narrative of text more often, and had more and better organized logic. Following the chronicity metaphor of expertise, Narvaez et al. (2006) showed that moral “chronics” respond faster and make more dispositional inferences when making judgments about morally relevant stories.²⁶ Similarly, people with higher moral reasoning scores showed a better recall of moral arguments in texts, and particularly better recall of more complex moral arguments (Narvaez & Gleason, 1995). This research program makes the case that at least some aspects of moral skill and knowledge follow the patterns one would expect from the expertise literature. And the use of expertise models in applied domains with clear moral relevance (Dane, 2010; Moulton et al., 2007) shows that others find the model a useful way to think about morally relevant action.

Box 6.1 shows the model of “mature moral functioning” Narvaez (2010) has developed from her research program. The complete model of moral expertise depicts moral experts as different in perception, cognition, action, metacognition, and meta-action. This kind of comprehensive expertise is based in extensive practice within a domain of morality (e.g. within a profession or calling). This facilitates the expert’s ability to be a more sensitive perceiver of ethically relevant occurrences, a more creative proposer of action plans, a better implementer of those plans, and a more thoughtful planner and networker in shaping the moral ecology.

This expansive, and even ideal, model of moral expertise is still limited by the recognition that moral expertise is typically developed within a domain. Even these highly skilled moral experts are limited in their expertise to those aspects of their world to which they have dedicated themselves. The work on moral exemplars seems to support this segmentation of moral expertise. People recognized for high moral achievement in social service or in a profession are often not such high achievers in other domains of their lives (e.g. their families are regularly negatively affected). It is unclear whether this neglect comes from inevitable competition for limited resources, lack of ethical commitment to domains other than the chosen one, or lack of some skill in the less well-practiced domain (Lapsley & Narvaez, 2006, p. 253).

One difficulty of the model of expertise provided here is that at times it seems to locate the action of expertise almost entirely in automatized responses to the

²⁶The term “chronics” is taken from the social cognition literature to refer to individuals whose mental models for a particular domain are most complex, detailed, and *chronically* accessible (Higgins et al., 1982).

Box 6.1 According to Narvaez (2010, p. 172), *mature moral functioning* involves:

- Basic socialization generally expected of adults (e.g. emotion regulation)
- Basic habits and disposition conducive to self-development
- Moral imagination
- Ethical expertise in a particular domain (e.g. a profession, or community service). Including
 - greater skills in ethical sensitivity
 - ethical judgment
 - ethical focus
 - flexible adaptation within networks of relationships
 - *individual* capacities for:
 - habituated empathic concern (targeted feelings of compassion combined with a sense of responsibility and propensity to act)
 - moral metacognition (manage and complete tasks, monitor progress, adapt plans and strategy), including moral locus of control, moral self-monitoring, and moral self-reflection
 - *collective* capacities for:
 - moral dialogue, commitment to and skill at initiating and supporting community dialogue about moral action
 - moral institutions, the construction, maintenance and reform of social institutions

This expertise looks different in every individual, and any particular moral expert cannot serve as the model for how all should operate. Note also that Narvaez (2010) conceives of moral expertise as occurring within domains, rather than being a more general, cross-domain expertise.

environment.²⁷ This may be because the model is based on the work on expertise by Dreyfus & Dreyfus (2004). Their model sees expertise as consisting primarily of highly practiced perception–judgment–action routines that develop over time. In the Dreyfus & Dreyfus (2004) model, these routines are seen as flexible and working under most circumstances. But what happens when they do not work?²⁸ We will look in detail at this issue in Section 6.4. For now we should note that the ability to be aware of the situation and to recognize when ordinary, practiced, expert routines are not working is itself a part of expertise, as is the ability to be reflective about the breakdown and to do appropriate puzzle-solving when normal solutions do not work (Moulton et al., 2007).

²⁷Though Narvaez (2010) is very explicit about the roles for thoughtful, conscious guidance of action.

²⁸Dreyfus and Dreyfus (2004, p. 256) do note that “in familiar but problematic situations, the expert deliberates about the appropriateness of his or her intuitions.” They do not give much explanation about how this deliberation works but they admit that “expert deliberation is not inferior to intuition” (p. 255). Still, to say that it is “not inferior” is small praise.

6.3.1 The Development of Expertise

Dreyfus and Dreyfus (2004) propose a five-stage model of skill acquisition. They begin their analysis with two morally neutral skills, driving and chess, and go on to consider the pattern of acquisition in the domain of morality.²⁹ The five steps along the continuum of expertise they describe are:

1. *Novice*: action is guided by context-free situational features and rules for determining actions based on those features of the situation. This simplistic approach fails in complex contexts.
2. *Advanced beginner*: guidance for action begins to be based in recognition of aspects of the context. At this level, maxims (e.g. make things as simple as possible) replace context-free rules.
3. *Competence*: the multiplicity of features and aspects that one must consider, even given maxims to help sort them, begins to overwhelm judgment because of combinatorial explosion. Hierarchical structures based on an individual's goals for that situation are imposed to limit the complexity. Goal choice is emotionally fraught because there is a lack of clear rules at this meta level.
4. *Proficiency*: action is guided by the adoption of a perspective that guides choices among higher-level goals. Flow in the task become continuous but is still punctuated by conscious planning and effort.
5. *Expertise*: "Normally an expert does not deliberate. He or she neither reasons nor acts deliberately. He or she simply spontaneously does what has normally worked and ... it normally works" (Dreyfus & Dreyfus, 2004, p. 253, ellipsis in original).

The stages are not discontinuous, and in appropriate environments, this development of expertise proceeds in tandem with appropriate practice. One can, for instance, see advances in expertise in the processing of moral discourse develop across childhood (Eisenberg, 2005; Narvaez & Gleason, 1995). *Kind* environments (Hogarth, 2001) provide support for the development of such expertise. Weather forecasting and sports are such environments, providing unambiguous, rapid feedback of success or error, and giving rise to internally consistent and externally valid mental models and intuition. When coupled with coaching in both practice and theory, these environments can support rapid expertise development. *Wicked* environments, on the other hand, do not provide the feedback needed because, for instance, definitions of success are malleable and feedback is ambiguous, difficult to recognize, or easy to reinterpret.³⁰ This makes it difficult to get (and easy to avoid) the feedback that can inform expert performance.

Given a kind environment, deliberate practice is necessary but not sufficient to achieve expert level performance. Campitelli and Gobet (2011) suggest it requires approximately 3,000 hours of deliberate practice, but this number varies tremendously (e.g. from hundreds to more than 20,000 hours to achieve master levels in

²⁹ But virtue ethics suggests that even these skills in a domain of practice are not morally neutral.

³⁰ This leads to the interesting and ironic hypothesis that the domain of morality may be wicked in this technical sense.

chess). This wide variability may be explained by individual differences in ability, sensitive periods for development (Campitelli & Gobet, 2011), and coaching (Hogarth, 2001). These numbers are likely to also vary widely by domain, and expertise may well not be a linear function of hours of deliberate practice (Campitelli & Gobet, 2011). One skill might be easy to learn at rudimentary levels (e.g. English) but difficult to excel at, while another may pose difficulties from the very beginning (e.g. Chinese). Where does morality stand in this? We do not know. The idea of morality as an expertise is still too young for these questions to have been investigated.

6.3.2 Automaticity in Expertise

There appears to be an odd irony in the Dreyfus and Dreyfus (2004) presentation of this progression in expertise. At the highest levels, automaticity rules performance, and the expert seems to do best when thinking least. This is a provocative idea, but there is some evidence to support it. Hassin and Bargh (2009) provide evidence that automatic processes do not need to be rigid and inflexible. Automatic-level goals can facilitate flexible action to achieve them, and automatic-level routines can even help resolve self-regulation dilemmas. One advantage of the automatic level of processing and action guidance is that it uses less energy, and thus frees the actor to concentrate on other things (Schmeichel, Vohs, & Baumeister, 2003). There is still controversy on the ways that goals and automated skills interact (Kruglanski & Szumowska, 2020; Mamede et al., 2010; Wood, Mazar, & Neal, 2021) but it lies with an agreement that a goal-flexible, adaptive automaticity could underlie expert negotiation of difficult tasks (and thus difficult moral tasks). In Section 6.4 we will look at another model of expertise that makes knowing when to drop out of automatic processing into deliberative processing a central aspect of expertise.

6.3.3 What Knowledge and Skill?

What kinds of skills and knowledge does one learn as one becomes more morally expert? The traditional Piagetian approach (Hatano & Inagaki, 2000; Narvaez, 2005) saw advancement driven by cognitive capability in foundational logical-mathematical operations. Advance in these operations supposedly produced advances across all domains of thought and action. This hierarchical structure is no longer viewed as valid in most circles (Hatano & Inagaki, 2000; Lapsley & Narvaez, 2006).³¹ Most work on conceptual development now sees it as similar to expertise in a domain, within the multiple domains of knowledge and action in our lives.³² So, the first broad answer to the question is that it will be knowledge and skill relevant to the domain of morality, and perhaps with regard to morality within a specific domain of action, such as a profession or one's family of origin. This does not seem like a remarkable statement unless one reflects on the short history of developmental theory outlined here.

³¹ This change is one reason behind the move to the “neo-Kohlberg” theory that depends on concepts like schemas and expertise (Narvaez, 2005).

³² We should note that there is some dispute among philosophers about whether knowledge is a necessary component of a virtue (Winter, 2011).

Beyond this quite vague description, can we specify what kind of knowledge is involved in expertise? Narvaez (2005) uses schema theory taken from Marshall (1995) to construct a model of moral skills and knowledge for the purpose of teaching them. She lists four types of knowledge:

1. *Identification*: knowledge of essential elements of the domain so that one can identify them.
2. *Elaboration*: knowledge that enables creation of a situational mental model.
3. *Planning*: knowledge of how to combine what is known to form expectations and make plans.
4. *Execution*: knowledge to guide the execution of plans and their revision when things go wrong.

Can one obtain reasonably valid descriptions of the relevant skills and knowledge in a domain? Much research in organizational psychology shows we can get reasonably valid descriptions of the skills involved in particular jobs (Fleishman & Mumford, 1991; Fleishman et al., 1991). In the domain of morality, attempts have been made to catalog and validate skills involved in moral judgment (Wainryb & Brehl, 2006) and moral behavior among school children (Narvaez, 2006). Lists of moral skills and knowledge have been developed for various professions, including computing professionals (Huff & Martin, 1995), dentistry (Bebeau, 1994), and medicine (Pinijphon, 2009).

How do these skills and knowledge differ between novices and experts? Keefer and Ashley (2001) asked experts in engineering ethics and undergraduate students in engineering to write responses to complex cases in engineering ethics. They found that expert ethical thinkers in this domain were more likely to:

- appeal to middle-level principles that identify role-specific obligations (e.g. due diligence in safety inspections)
- make greater use of professional knowledge in order to recognize moral issues and relevant facts (e.g. knowledge about interactions among safety aspects of particular systems)
- employ more contextually sensitive reasoning strategies when crafting resolutions including:
 - identifying alternative moral issues
 - assessing the moral implications of actions
 - providing “alternative practical resolutions”

One surprising aspect of the findings is that novices in the domain tended to use the more abstract, “philosophical” reasoning strategies, while domain experts (many of whom were philosophers) used middle-level concepts and role-specific obligation to identify problems and think about the cases.

The discussion so far makes it seem as though the skills and knowledge outlined here are entirely the work of the isolated individual. But at least some of the skill one has in identifying moral issues comes from knowing how to interact with others to investigate and understand their perspectives. Similar social skills can be found in each of the categories listed earlier. Chapter 3 makes this point in detail.

6.3.4 Domain Specificity

We have already noted that expertise is domain specific. This is in part what one *means* when one speaks of expertise – expertise is always expertise-in-something. The virtues that moral exemplars in engineering listed for other engineers include more general items such as integrity and courage, but also domain-specific items like “habit of documenting work” and “competence” (Pritchard, 1998, 2001). One of the computing moral exemplars that one author (CH) interviewed commented that technical and creative *in*competence can produce additional ethical temptation to cut corners and to do sub-par, unethical work (Huff & Barnard, 2009).

Others have noted that moral exemplars in social service often find themselves regretting the ways they have treated their families. Colby and Damon (1992) speculate that this is not a simple case of not having concern for their families, but more likely a result of conflict among moral goals given limited resources. They base this hypothesis on the ways their exemplars spoke about valuing their families’ welfare, feeling compelled to spend time on their other moral commitments, and feeling guilt and distress over the neglect this produced. Given the inevitable limited resources individuals have, it seems reasonable there might actually be competition among domains (Lapsley & Narvaez, 2006). So, at least in this case, it may not be a lack of skill in the domain of the family that produces the neglect. But if resource limitations lead to less practice in one domain (e.g. family) it may lead to less knowledge about that domain and less skill in practicing in that domain. Thus, those led to neglect a domain may find themselves less able to practice in it when they do finally find the time; and they may find their well-honed skills in other domains less successful than they might hope.

6.4 Habits

When one speaks of habits, the term is often modified by an implied “bad.” But in his *Nicomachean Ethics*, Aristotle suggests that the virtuous person was prepared (or “cultivated”) for virtue by learning good habits (Aristotle, 340 BCE/1941, NE 10.9 1179b4-31). The importance of habits for moral behavior can be seen from how much of our overall behavior they control. In experience sampling studies, habitual actions comprise 45% of everyday activity (Neal, Wood, & Quinn, 2006).

In the psychological literature, habits are “automated response dispositions that are cued by aspects of the performance context” (Neal et al., 2006, p. 198) and in terms of value may be good, bad, or indifferent. Habits can be simple aspects of our daily routine (e.g. how we squeeze the toothpaste tube) or they can be relatively complex parts of an expert routine (e.g. an engineer checking for safety in the interactions of components). But to count as habitual, they must operate primarily automatically and be controlled by cues in the individual’s context. They are learned by repetition in a similar context so that procedural memory – that aspect of memory that stores and controls action sequences – is “cognitively tuned” to both the context and the response. This suggests their relationship to expertise: those parts of expertise that can be described as “automated response dispositions” can be thought of as habits. But as we will see, expertise is more than a collection of habits.

Still, even habits are not simple chains of stimulus–response links without representations of the environment or goals. There are roles for consciousness, goals, and motivational states in habits (Neal et al., 2006). At the simple habit level, Rothman, Sheeran, and Wood (2009) propose a model of habit change that includes both automatic and controlled processes that regulate eating behavior at different times in the behavior change process.

Some habits are controlled almost entirely by performance context cues. Others are not always performed in a particular context but their performance can be triggered, without conscious intervention, by particular goals in the appropriate context. The more a habit is under the control of environmental cues, and the more highly practiced it is, the more difficult it is to intentionally alter (Neal, Wood, Labrecque, & Lally, 2012).³³ When a particular action only occurs rarely, or is performed in complex interaction with goals and contexts, it may require a great deal of practice to become habitual (Neal et al., 2012). This helps to make sense of the long practice times required to cultivate expertise, which will consist of a multitude of habits that are integrated with both goals and contexts.

One advantage of the mostly automatic nature of habits is that they do not require much in the way of cognitive resources (Neal et al., 2006). In complex or stressful contexts, this leaves the actor with the cognitive resources to concentrate on other aspects of performance. Of course, if the habit is not appropriately sensitive to context changes, it may be performed at inappropriate times. Expertise may consist in part of tuning habits so that they are expressed only in the appropriate contexts. The lack of such tuning is called “cognitive entrenchment” in the organizational literature (Dane, 2010).³⁴

6.4.1 How Are Habits Regulated?

The appropriate integration of habits into expertise requires that they be sensitive to goals and contexts, even if they are so well learned that they are automatic in activation and execution. Neal, Wood, and Quin (2006, p. 200) call this process of integration

³³In the habit literature, habits that are always performed in a particular context are termed “strong” habits while those that are deployed depending on the goal and the context are called “moderate strength” habits. This language is one of the last vestiges of the behaviorist interest in habits, and it can be misleading. It considers habits as isolated units that are stronger or weaker, rather than as possibly integrated units in a larger expertise in a domain. One might easily call habits that are activated only in the appropriate context “flexible habits,” or even “strong habits,” because they are tightly integrated into expertise.

³⁴This detachment of habits from context has the ironic effect of making habits feel unconnected to the sense of self. Thus, people think their non-habitual actions are more informative about who they are (Neal et al., 2006, p. 201). This is perhaps because the self is autobiographical and more dependent on declarative, narrative memory. This disconnection of habit from the self could be one reason for the “humility” shown in exemplars. It might also help us understand why we think our judgment in difficult (non-automatic) situations is the marker of our moral character rather than the developed automatized skills of moral expertise.

self-regulation,³⁵ and note that it consists of “monitoring and adjusting responses in the service of the self.”

But if habits are automatic, how can this kind of self-regulation interact with them? One suggestion is that the perception of some metacognitive difficulty or disfluency in the process is a cue for conscious intervention (Alter et al., 2007; Moulton et al., 2007; Neal et al., 2006). For instance, when an individual has multiple goals, these can conflict with each other and activate competing responses. When complex contexts activate competing habits, this too may cause disfluency. Or when a goal is blocked (perhaps by some conflict with the context, such as a pattern of data that does not fit with a medical diagnosis) disfluency may be experienced.

Thus, experts need to use markers of disfluency (Moulton et al., 2007) to break out of automatic routines and respond consciously to these difficulties. Moulton et al. (2007) have a model of “slowing down when you should” and their model of expertise showcases the ability to switch back and forth from automatic to controlled processes. Their work is based in research on medical expertise, and they draw from a model of expertise by Schön (1984) taken from engineering, architecture, management, psychotherapy, and town planning. In contrast to the Dreyfus and Dreyfus (2004) model mentioned earlier of “doing what normally works,” they ask how expertise copes in extraordinary circumstances. They mark the ability to switch back and forth from automatic to controlled process as the critical skill needed for expertise in extraordinary circumstances. This ability to self-regulate is facilitated by awareness of the situation that allows detection of disfluency, the connection of that disfluency to situational parameters, and the initiation of appropriate corrective action (Alter et al., 2007). The expert then has a well-practiced controlled process to search for alternative solutions to the issue, both naming the difficulty and framing it, placing it in the larger picture of the situation and goals (Schön, 1984).³⁶ Some of the psychological processes underlying this flexibility of automatic action have been described by De Neys (2012, 2021), though again there is disagreement over how the processes play out (Kruglanski & Szumowska, 2020; Mamede et al., 2010; Wood et al., 2021).

Another sort of disfluency can come from recognition that one’s performance in different domains is differentially successful. One barrier to the sort of integrity that is achieved from cross-domain consistency is competition between domains. Competition can arise from simple conflicts for limited resources (e.g. time) or from more complex conflicts when the goals associated with different cultural or personal domains are in conflict (Schwartz, 2010). When we speak of moral integrity (Mumford et al., 2001; Paine, 1994; Pritchard, 2006), we often refer to consistency in value commitments or consistency of behavior with a particular value commitment (see Chapter 5). But if we think of moral action in the context of domain dependent expertise, it becomes clear that one might have commitment to a domain but not the requisite level of expertise in it. And one can thus speak of

³⁵Note that this is a different use of the term than what we have used in Section 6.1.2. It is not unusual for such terms to be treated slightly differently by different authors (see, e.g. the table of definitions for moral psychology in the Introduction).

³⁶This process is reminiscent of the hermeneutic framework that Gadamer (1996) provides in his work on medical practice.

lack of skill as a threat to integrity. Integrity can be threatened by the inevitable competition for limited resources, or from a lack of ethical commitment to domains other than the chosen one, or from a lack of some skill in a less well-practiced domain. Restoring integrity, then, involves the skill of recognizing the disfluency and moving to address it.

The expert has the cognitive capacity to be aware of disfluency and situational context in part because the automatized habits require little capacity to run. Habits are thus a two edged sword: they can both free the expert to respond flexibly to the extraordinary (Moulton et al., 2007) and trap the expert into routines that are not appropriate to the context (Dane, 2010).

6.5 Skills Go Awry

Finally, we must make a comment about the *moral* in the phrase moral skill. What is peculiarly moral about these skills? Psychopaths can exhibit significant self-regulation skills in the service of their goals, but still show “feckless disregard” and meanness toward others (Skeem et al., 2011, p. 107). This seems to be based in the facet of psychopathy associated with sensitivity to norms (Luke, Neumann, & Gawronski, 2021). The classic work on German police participation in the final solution describes at least one subset of the most efficient “Jew-hunters” as having particular emotional regulation skills that allowed them to distance themselves from their work (Browning, 1992, p. 127) and thus to participate willingly in mass murder.

Thus, the relationship between morality and the skills and expertise we outline here is at the very least complex, if not contradictory and dialectical.³⁷ Skills and knowledge can be separated from morality and can indeed serve immoral ends. As we have seen in the various models of moral exemplars (e.g. Colby & Damon, 1992; Walker & Hennig, 2004), moral action and its influences are characterized by equifinality – the same moral action can come about in many different ways. And any particular influence (e.g. a skill, a personality characteristic, etc.) is multi-final – it can support or contribute to a range of both moral and immoral action. But in most instances, absent moral luck (Nagel, 1979; Williams, 1981), moral actors must have some skills to successfully obtain their goals. Thus, if we want to understand, and ultimately to help develop, successful moral action, we will want to understand the role that skill plays in it.

A final thing to note regarding the moral reach of skills is their seeming domain specificity. Even more general skills, such as emotion regulation, can manifest differently in different life domains (Harley et al., 2019). And one can easily imagine that things like moral attentiveness, the ability to recognize moral issues, would depend on knowledge and skill in the particular domain in question (e.g. Nathaniel Borenstein’s recognition of the dangers of particular military software).

³⁷ See the discussion on this in Chapter 9.

6.6 Discussion

When we include taking action in the domain of morality, expertise based in knowledge, skill, and habit becomes immediately relevant. This approach is both a description of how one operates in a domain and a description of an ideal state toward which one constantly strives. How might this change our picture of moral action?

6.6.1 Conclusion

1. *Principled decision making is likely not the primary path to taking moral action.*

Moral judgment and moral decision-making may be important skills when operating in difficult situations, but much skilled moral action is not consciously guided. From the perspective of skill, expertise, and habit, conscious moral judgment and decision-making of the sort that is measured in moral reasoning tests is likely an exceptional occurrence. Moral acting in everyday life may more often consist of automatically or habitually pursuing a moral goal rather than constantly deliberating about moral dilemmas.

Principled decision-making may not be primary in terms of being learned first, or in terms of being foundational. It is probably one necessary but not sufficient ingredient in the ongoing behavior stream of expertise. It may be most likely to be called upon when some disfluency occurs in action that needs resolution, but even then it may involve intermediate level knowledge in a domain rather than any return to foundational ethical principles. The only way you can claim that principled decision-making is the primary path to moral action is to *define* moral action as action that occurs based on principled decision-making. And to do this leaves out a great deal of moral action.

2. *Many of the skills and habits that support moral action are domain limited.*

Expertise is always expertise in some domain. To discover this, one needs only to venture outside that domain (e.g. attempt casual conversation or humanitarian assistance in another culture). Noticing and responding to a moral issue in one's family requires different skills and knowledge than doing so at work. Moral exemplars in social service often find themselves regretting the ways they have treated their families. There are more general level skills, such as emotion regulation, that underlie moral action across domains, but even these are likely to have domain limitations. It may well be that this domain specificity of skills contributes to the problem of behavioral integrity that we review in Chapter 5.

3. *By designing kind environments, we can leverage the learning of moral expertise in a domain.*

If we treat moral action as a design problem (see Chapter 3), we can better support the acquisition of moral expertise in new domains. We can identify and teach skills that are automatized in moral action and problem-solving skills that operate consciously when disfluency in automatic action requires it. Concentrating exclusively on moral judgment will not be enough to support expertise in moral action. When possible, we can provide unambiguous, rapid feedback that supports the development of internally consistent and externally valid mental models and intuition that can underlie automated moral action. Some domains will be more conducive to this approach than others. Value conflicts in some domains will make it harder, and some disfluency in moral action may be due to these conflicts (rather than character).

6.6.2 Application

On Saturday April 30, 1977, fourteen women gathered for the first demonstration of the Mothers of the Plaza de Mayo. They ignored the warnings of danger from their families and gathered at the center of political power in Argentina to give witness to the disappearances and demand answers. Some of them were themselves later disappeared. These women, and the thousands who joined them, are not famous for their good will. They are famous for their fierce determination, their dogged resistance, their courage, and their skill in resistance and public protest. They are in short famous for courageous skilled moral action.

The original members met each other while visiting the same bureaucratic offices in search of their children. They banded together to protest, then organized to march, and then initiated an international information campaign to counter government propaganda. They knew how to do some of these things, they also knew how to find people with additional skills and how to convince them to take the risk of joining the movement. Many are still marching today and running websites and international campaigns to keep up the search for those still lost and to prevent the current government from reducing penalties on those convicted of crimes.

6.6.3 Open Questions

The idea of moral skill has been embedded in virtue models of moral action since before Aristotle and Confucius. But there has been remarkably little empirical work on the skilled action aspects of moral action. We are still in the beginning stages:

1. *What skills are involved in what cultures and domains of moral action?*

There are numerous applied attempts to catalog the skills involved in moral action, but these all suffer from the same “laundry list” complaints made about virtue approaches (see Chapter 4). Different lists often have little overlap, and the inclusion or exclusion of particular skills seems arbitrary and unrelated to systematic theory. Given the domain-based nature of expertise, one should not expect a complete skill set that is applicable across all domains. Instead, one might expect some core skills and a large proliferation of domain-specific skills in the various cultures, professions, vocations, and contexts in which one can take moral action. In parallel with the literature on leadership, these contexts may interact in complex ways with characteristics of the actor and moral ecology. For this reason, work in this area might best start with careful empirical attention to domain and cultural differences and then search for shared or overlapping items. One can then ask which skills might be transferrable across which domain boundaries.

2. *How do actors negotiate the interplay between controlled and automatic moral action?*

Increases in expertise move moral action from the controlled to the automatic. High levels of expertise require subtle shifts between automatic and controlled processing and action. We are only beginning to understand the integration of automatic and controlled action, much less moral action. What controls the interplay between these two ways of taking moral action? What environmental and personal characteristics make cognitive entrenchment more or less likely? What

skills are required to recognize disfluency and take controlled action based on it? Are there different kinds of disfluency and do they interact with different characteristics of the environment (e.g. kind/wicked)?

3. *How do people understand and learn from moral failure?*

The point of the discussion in this chapter about disfluency is that one can (but often does not have to) recognize when one's skilled performance is failing. What are the characteristics of the actor and the situation that allow this recognition to occur and that facilitate effective reparative action? The list of skills associated with moral action present multiple points of failure in this process, including ones that may differ from one domain to another. Self-regulation involves the organization of attentional, motivational, emotional, and other internal resources in the service of goals. There are multiple points of failure associated with each of these. One can fail in deciding which of many alternate goals to pursue, in planning how best to attain a valued end, in protecting goals from competing concerns, in recovery when initial efforts fail, and in deciding whether to continue or abandon goals after failure. The resilience of each of these failure points may also differ depending on the surrounding moral ecology.

4. *How does the expertise model help us understand cultural disagreement about moral action?*

One way that cultural disagreements can be viewed is as differences in expertise in different domains or moral ecologies. The knowledge and skill needed for expert-level moral acting in an individualistic culture may be quite different from that needed for a collectivist culture (not to mention more specific aspects of a culture, e.g. particular religious commitments). There is no requirement that experts always agree (otherwise it would be a simple rule-oriented domain and require no expertise). If we think about cultural disagreement as disagreement among experts with different kinds of expertise, we may find some guidance regarding the nature, structure, and extent of the moral disagreements one can find between cultures.

6.7 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Benita (2020) "Freedom to feel: A self-determination theory account of emotion regulation." A recent review of emotion regulation literature that focuses on the autonomy of the individual.
- Dreyfus and Dreyfus (2004) "The ethical implications of the five-stage skill-acquisition model." Experts in expertise apply their model to moral action.
- Koole (2009). "The psychology of emotion regulation: An integrative review." A model of emotional regulation that analyses it as the intersection of its functions (goal, need, and person) and the involved systems (attention, knowledge, and body).
- Kuhl et al. (2021). "The functional architecture of human motivation." The most recent summary of personality systems interactions theory that attempts to locate many personality skills in interdependent systems.

- Narvaez (2005). "The neo-Kohlbergian tradition and beyond: Schemas, expertise, and character." An early paper proposing skills and expertise that support moral action.
- Narvaez (2010) "Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning." A model of mature moral functioning looked at through the lens of expertise across a variety of systems (e.g. cognition, emotion, etc.).
- Wood et al. (2021) "Habits and goals in human behavior: Separate but interacting systems." A review of work on habits that concludes that habits can be activated independent of goals but they can also be integrated into goal pursuit.

References

- Alter, A. L., Oppenheimer, D. M., Epley, N. et al., (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, 136(4), 569–576. doi:10.1037/0096-3445.136.4.569
- Anderson, M. S., Horn, A. S., Risbey, K. R. et al., (2007). What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a national survey of NIH-funded scientists. *Academic Medicine*, 82(9), 853–860.
- Angus, L., Watson, J. C., Elliott, R. et al. (2015). Humanistic psychotherapy research 1990–2015: from methodological innovation to evidence-supported treatment outcomes and beyond. *Psychotherapy Research*, 25(3), 330–347. doi:10.1080/10503307.2014.989290
- Annas, J. (2006). Virtue ethics. In D. Copp (Ed.), *The Oxford Handbook of Ethical Theory* (pp. 515–536). New York: Oxford University Press.
- Anscombe, G. E. M. (1958). Modern Moral Philosophy. *Philosophy*, 33, 1–19.
- Aristotle. (340 BCE/1941). *Nicomachean ethics* (W. D. Ross, Trans.). In R. McKeon (Ed.), *The Basic Works of Aristotle* (pp. 927–1112). New York: Random House.
- Arnold, M. B. (1960). *Emotion and Personality*. New York: Columbia University Press.
- Baumeister, R. F., Gailliot, M. T., & Tice, D. M. (2009). Free willpower: A limited resource theory of volition, choice, and self-regulation. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford Handbook of Human Action* (pp. 487–508). Oxford: Oxford University Press.
- Baumeister, R. F., & Tierney, J. (2011). *Willpower: Rediscovering the Greatest Human Strength*. New York: Penguin Press HC.
- Bebeau, M. J. (1994). Influencing the moral dimensions of dental practice. In J. R. Rest & D. Narvaez (Eds.), *Moral Development in the Professions: Psychology and Applied Ethics* (pp. 121–146). Hillsdale, NJ: Lawrence Erlbaum.
- Benita, M. (2020). Freedom to feel: A self-determination theory account of emotion regulation. *Social and Personality Psychology Compass*, 14(11). doi:10.1111/spc3.12563
- Bernheim, B. D. R., Antonio. (2004). Addiction and cue-triggered decision processes. *The American Economic Review*, 94(5), 1558–1590.
- Boehm, B., & Turner, R. (2003). Using risk to balance agile and plan-driven methods. *Computer*, 36(6), 57–66. doi:10.1109/MC.2003.1204376
- Boekarts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2005). *Handbook of Self-Regulation*. San Diego: Elsevier.
- Bono, G., Emmons, R. A., & McCullough, M. E. (2012). Gratitude in practice and the practice of gratitude. In P. A. Linley & S. Joseph (Eds.), *Positive Psychology in Practice* (pp. 464–481). New York: Wiley.

- Breuer, J., & Freud, S. S. (1895/2000). *Studies in Hysteria* (J. Strachey & A. Freud, Trans.).
- Browning, C. R. (1992). *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York: HarperCollins.
- Callahan, D. (1980). Goals in the teaching of ethics. In D. Callahan & S. Bok (Eds.), *Teaching Ethics in Higher Education* (pp. 61–74). New York: Plenum.
- Campitelli, G., & Gobet, F. (2011). Deliberate practice: Necessary but not sufficient. *Current Directions in Psychological Science*, 20(5), 280–285. doi:10.1177/0963721411421922
- Carver, C. S., & Scheier, M., F. (2002). Control processes and self-organization as complimentary principles underlying behavior. *Personality and Social Psychology Review*, 6(4), 304–315.
- Cervone, D., Shadel, W. G., Smith, R. E. et al. (2006). Self-regulation: Reminders and suggestions from personality science. *Applied Psychology*, 55(3), 333–385.
- Colby, A., & Damon, W. (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. New York: Free Press.
- Condon, P., & Makransky, J. (2020). Recovering the relational starting point of compassion training: A foundation for sustainable and inclusive care. *Perspectives on Psychological Science*, 15(6), 1346–1362. doi:10.1177/1745691620922200
- Crowell, C. R., Narvaez, D., & Gombert, A. (2004). Moral psychology and information ethics. In L. Freeman & G. Peace (Eds.), *Information Ethics: Privacy and Intellectual Property* (pp. 19–37). Hershey, PA: Information Science Publishing.
- Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *Academy of Management Review*, 35(4), 579–603.
- Davenport, J. (2007). *Will as Commitment and Resolve*. New York: Fordham University Press.
- De las Fuentes, C., Willmuth, M. E., & Yarrow, C. (2005). Competency training in ethics education and practice. *Professional Psychology: Research and Practice*, 36(4), 362–366.
- De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, 7(1), 28–38. doi:10.1177/1745691611429354
- De Neys, W. (2021). On dual- and single-process models of thinking. *Perspectives on Psychological Science*, 16(6), 1412–1427. doi:10.1177/1745691620964172
- de Ridder, D. T. D., Lensvelt-Mulders, G., Finkenauer, C. et al. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors. *Personality and Social Psychology Review*, 16(1), 76–99. doi:10.1177/1088868311418749
- de Waal, F. B. M. (2009). Darwin's last laugh. *Nature*, 460(7252), 175.
- Decety, J. (2021). Why empathy is not a reliable source of information in moral decision making. *Current Directions in Psychological Science*, 30(5), 425–430. doi:10.1177/096372142111031943
- Depow, G. J., Frances, Z., & Inzlicht, M. (2021). The experience of empathy in everyday life. *Psychological Science*, 32(8), 1198–1213. doi:10.1177/0956797621995202
- Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science*, 15(4), 942–956. doi:10.1177/1745691620904771
- Dreyfus, H. L., & Dreyfus, S. E. (2004). The ethical implications of the five-stage skill-acquisition model. *Bulletin of Science, Technology, & Society*, 24, 251–264. doi:10.1177/0270467604265023
- Duckworth, A. L., Peterson, C., Matthews, M. D. et al. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101. doi:10.1037/0022-3514.92.6.1087
- Eisenberg, N. (2005). The development of empathy-related responding. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (pp. 72–117). Lincoln, NE: University of Nebraska Press.
- Ent, M. R., Baumeister, R. F., & Tice, D. M. (2015). Trait self-control and the avoidance of temptation. *Personality and Individual Differences*, 74, 12–15. doi:10.1016/j.paid.2014.09.031

- Feinberg, M., Willer, R., Antonenko, O. et al. (2012). Liberating reason from the passions: Overriding intuitionist moral judgments through emotion reappraisal. *Psychological Science*, 23(7), 788–195. doi:10.1177/0956797611434747
- Fleishman, E. A., & Mumford, M. D. (1991). Evaluating classifications of job behavior: A construct validation of the ability requirement scales. *Personnel Psychology*, 44(3), 523–575.
- Fleishman, E. A., Mumford, M. D., Zaccaro, S. J. et al. (1991). Taxonomic efforts in the description of leader behavior: A synthesis and functional interpretation. *The Leadership Quarterly. Special Issue: Individual Differences and Leadership: I*, 2(4), 245–287. doi:10.1016/1048-9843(91)90016-u
- Förster, J., & Jostmann, N. B. (2012). What is automatic self-regulation? *Zeitschrift für Psychologie*, 220(3), 147–156. doi:10.1027/2151-2604/a000107
- Fowler, Z., Law, K. F., & Gaesser, B. (2021). Against empathy bias: The moral value of equitable empathy. *Psychological Science*, 32(5), 766–779. doi:10.1177/0956797620979965
- Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. Cambridge, MA: MIT Press.
- Friese, M., Loschelder, D. D., Gieseler, K. et al. (2019). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, 23(2), 107–131. doi:10.1177/1088868318762183
- Frijda, N. H., & Sundararajan, L. (2007). Emotion refinement: A theory inspired by Chinese poetics. *Perspectives on Psychological Science*, 2(3), 227–241.
- Frimer, J. A., & Walker, L. J. (2008). Towards a new paradigm of moral personhood. *Journal of Moral Education. Special Issue: Towards an Integrated Model of Moral Functioning*, 37(3), 333–356. doi:10.1080/03057240802227494
- Gadamer, H.-G. (1996). *The Enigma of Health: The Art of Healing in a Scientific Age*. Stanford, CA: Stanford University Press.
- Galotti, K. M. (2007). Decision structuring in important real-life choices. *Psychological Science*, 18(4), 320–325.
- Garland, E. L., Farb, N. A., Goldin, P. R. et al. (2015). Mindfulness broadens awareness and builds eudaimonic meaning: A process model of mindful positive emotion regulation. *Psychological Inquiry*, 26(4), 293–314. doi:10.1080/1047840x.2015.1064294
- Goldenberg, A., Garcia, D., Halperin, E. et al. (2020). Collective emotions. *Current Directions in Psychological Science*, 29(2), 154–160. doi:10.1177/0963721420901574
- Gollwitzer, P. M., Sheeran, P., Trotschel, R. et al. (2011). Self-regulation of priming effects on Behavior. *Psychological Science*, 22(7), 901–907 doi:10.1177/0956797611411586
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101–124. doi:10.1080/1047840x.2012.651387
- Hannah, S., T., & Avolio, B. J. (2010). Moral Potency: Building the capacity for character-based leadership. *Consulting Psychology Journal*, 62(4), 291–310.
- Harley, J. M., Pekrun, R., Taxer, J. et al. (2019). Emotion regulation in achievement situations: An integrated model. *Educational Psychologist*, 54(10), 1–21. doi:10.1080/00461520.2019.1587297
- Hassin, R. R., Bargh, J. A., & Zimmerman, S. (2009). Automatic and flexible: The case of nonconscious goal pursuit. *Social Cognition*, 27(1), 20–36.
- Hatano, G., & Inagaki, K. (2000). Domain specific constraints of conceptual development. *International Journal of Behavioral Development*, 24(3), 267–275.
- Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, 43(1), 35–47.
- Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, 4(2), 162–176.
- Hogarth, R. M. (2001). *Educating Intuition*. Chicago: University of Chicago Press.

- Hölzel, B. K., Lazar, S. W., Gard, T. et al. (2011). How does mindfulness meditation work? Proposing mechanism of action from a conceptual and neural perspective. *Perspectives on Psychological Science*, 6(6), 537–559. doi:10.1177/1745691611419671
- Hudac, C. M., & Sommerville, J. A. (2020). Understanding other's minds and morals: Progress and innovation of infant electrophysiology. In J. Decety (Ed.), *The Social Brain: A Developmental Perspective* (pp. 179–197). Cambridge, MA: MIT Press.
- Huff, C., & Barnard, L. K. (2009). Good computing: moral exemplars in the computing profession. *IEEE Technology and Society Magazine* (Fall), 47–54.
- Huff, C., & Martin, C. D. (1995). Computing consequences: A framework for teaching ethical computing. *Communications of the ACM*, 38(12), 75–84.
- Hulsey, T. L., & Hampson, P. J. (2014). Moral expertise. *New Ideas in Psychology*, 34(0), 1–11. doi:http://dx.doi.org/10.1016/j.newideapsych.2014.02.001
- Kant, I. (1785/2011). *Groundwork of the Metaphysics of Morals: A German–English Edition* (M. Gregor & J. Timmermann, Trans.). Cambridge: Cambridge University Press.
- Keane, W. (2015). *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton University Press.
- Keefer, M. W., & Ashley, K. D. (2001). Case-based approaches to professional ethics: A systematic comparison of students' and ethicists' moral reasoning. *Journal of Moral Education*, 30(4), 377–398.
- Kierkegaard, S. (1849/2004). *The Sickness Unto Death* (A. Hannay, Trans.). New York: Penguin.
- Kok, B. E., & Singer, T. (2017). Phenomenological fingerprints of four meditations: Differential state changes in affect, mind-wandering, meta-cognition, and interoception before and after daily practice across 9 months of training. *Mindfulness (N Y)*, 8(1), 218–231. doi:10.1007/s12671-016-0594-9
- Koole, S. L. (2009). The psychology of emotion regulation: An integrative review. *Cognition & Emotion*, 23(1), 4–41. doi:10.1080/02699930802619031
- Koole, S. L., McCullough, M. E., Kuhl, J. et al. (2010). Why religion's burdens are light: From religiosity to implicit self-regulation. *Personality and Social Psychology Review*, 14(1), 95–107. doi:10.1177/1088868309351109
- Kruglanski, A. W., & Szumowska, E. (2020). Habitual behavior is goal driven. *Perspectives on Psychological Science*, 15(5), 1256–1271. doi:10.1177/1745691620917676
- Kuhl, J., & Koole, S. L. (2004). Workings of the will: A functional approach. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of Experimental Existential Psychology* (pp. 411–430). New York: Guilford Press.
- Kuhl, J., Quirin, M., & Koole, S. L. (2021). The functional architecture of human motivation: Personality systems interactions theory. In A. Elliot (Ed.), *Advances in Motivation Science* (Vol. 8, pp. 1–62). New York: Elsevier.
- Lapsley, D. K. (2008). Moral Self-Identity as the aim of education. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of Moral and Character Education* (pp. 30–52). New York: Routledge.
- Lapsley, D. K., & Narvaez, D. (2005). Moral psychology at the crossroads. In D. K. Lapsley & F. C. Power (Eds.), *Character Psychology and Character Education* (pp. 18–35). South Bend, IN: University of Notre Dame Press.
- Lapsley, D. K., & Narvaez, D. (2006). Character education. In W. Damon, R. M. Lerner, A. K. Renninger et al. (Eds.), *Handbook of Child Psychology: Child Psychology in Practice* (Vol. 4, pp. 248–296). New York: Wiley.
- Latane, B., & Darley, J. M. (1970). *The Unresponsive Bystander: Why Doesn't He Help?* Upper Saddle River, NJ: Prentice Hall.
- Lin, P., Saunders, B., Friese, M. et al. (2020). Strong effort manipulations reduce response caution: a preregistered reinvention of the ego-depletion paradigm. *Psychological Science*, 31(5), 531–547. doi:10.1177/0956797619893078

- Lord, R. G., Diefendorff, J. M., Schmidt, A. M. et al. (2009). Self-regulation at work. *Annual Review of Psychology*, 61, 543–568. doi:10.1146/annurev.psych.093008.100314
- Ludwig, V., U. Brown, K. W., & Brewer, J. A. (2020). Self-regulation without force: Can awareness leverage reward to drive behavior change? *Perspectives on Psychological Science*, 15(6), 1382–1399. doi:10.1177/1745691620931460
- Luke, D., Neumann, C., & Gawronski, B. (2021). Psychopathy and moral-dilemma judgment: An analysis using the four-factor model of psychopathy and the CNI model of moral decision making. *Clinical Psychological Science*, 10(3), 553–569. doi:10.1177/21677026211043862
- MacIntyre, A. C. (1981). *After Virtue: A Study in Moral Theory* (Vol. 1985). London: Duckworth.
- Mamede, S., Schmidt, H. G., Rikers, R. M. et al. (2010). Conscious thought beats deliberation without attention in diagnostic decision-making: at least when you are an expert. *Psychological Research*, 74(6), 586–592. doi:10.1007/s00426-010-0281-8
- Marshall, S. P. (1995). *Schemas in Problem Solving*. New York: Cambridge University Press.
- McGregor, I., & Little, B. R. (1998). Personal projects, happiness, and meaning: On doing well and being yourself. *Journal of Personality and Social Psychology*, 74(2), 494–512.
- Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, 244(4907), 933–938. doi:10.1126/science.2658056
- Morris, A., Phillips, J., Huang, K. et al. (2021). Generating options and choosing between them depend on distinct forms of value representation. *Psychological Science*, 32(11), 1731–1746. doi:10.1177/09567976211005702
- Motokio, Z. (1400/1984). On the art of the Nō drama: The major treatises of Zeami. In M. Yamazaki (Ed.). Princeton, NJ: Princeton University Press.
- Moulton, C.-a. E., Regehr, G., Mylopoulos, M. et al. (2007). Slowing down when you should: A new model of expert judgment. *Academic Medicine*, 82(10), S109–S116.
- Mumford, M. D., Connelly, M. S., Helton, W. et al. (2001). On the construct validity of integrity tests: Individual and situational factors as predictors of test performance. *International Journal of Selection and Assessment*, 9(3), 240–257.
- Mumford, M. D., Connelly, S., Brown, R. P. et al. (2008). A sensemaking approach to ethics training for scientists: Preliminary evidence of training effectiveness. *Ethics & Behavior*, 18(4), 25p. doi:10.1080/10508420802487815
- Murata, K. (2010). Information ethics and culture: Revisiting Japanese ethical values. Paper presented at the ETHICOMP2010, Barcelona, Spain.
- Nagel, T. (1979). *Mortal Questions*. New York: Cambridge University Press.
- Narvaez, D. (1998). The influence of moral schemas on the reconstruction of moral narratives in eighth graders and college students. *Journal of Educational Psychology*, 90(1), 13–24. doi:10.1016/S0022-5371(79)90146-4
- Narvaez, D. (2005). The neo-Kohlbergian tradition and beyond: Schemas, expertise, and character. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol. 51: Moral Motivation Through the Life Span* (pp. 119–163). Lincoln, NE: University of Nebraska Press.
- Narvaez, D. (2006). Integrative ethical education. In M. Killen & J. Smetana (Eds.), *Handbook of Moral Development* (pp. 703–733). Mahwah, NJ: Erlbaum.
- Narvaez, D. (2010). Moral Complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Narvaez, D., & Gleason, T. (1995). The relation of moral judgment development and educational experience to recall of moral narratives and expository texts. *The Journal of Genetic Psychology*, 156(1), 1–23.
- Narvaez, D., Lapsley, D. K., Hagele, S. et al. (2006). Moral chronicity and social information processing: Tests of a social cognitive approach to the moral personality. *Journal of Research in Personality*, 40(6), 966–985. doi:10.1016/j.jrp.2005.10.003

- Neal, D. T., Wood, W., Labrecque, J. S. et al. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *Journal of Experimental Social Psychology*, 48(2), 492–498. doi:10.1016/j.jesp.2011.10.011
- Neal, D. T., Wood, W., & Quinn, J. M. (2006). Habits – A repeat performance. *Current Directions in Psychological Science*, 15(4), 198–202.
- Norcross, J. C., Krebs, P. M., & Prochaska, J. O. (2011). Stages of change. *Journal of Clinical Psychology*, 67(2), 143–154. doi:10.1002/jclp.20758
- Nussbaum, M. C. (1996). *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*. Princeton, NJ: Princeton University Press.
- Opatow, S. (2010). Social injustice and moral exclusion. In N. J. Young (Ed.), *The Oxford International Encyclopedia of Peace*. New York: Oxford University Press.
- Paine, L. S. (1994). Managing for organizational integrity. *Harvard Business Review*, 72(2), 106–117.
- Piniiphon, P. (2009). An ICM approach to the assessment of a medical ethics intervention in Thailand. PhD dissertation. The University of Alabama, Tuscaloosa.
- Porat, R., Tamir, M., Halperin, E. et al. (2020). Group-based emotion regulation: A motivated approach. *Emotion*, 20(1), 16–20.
- Preti, A. (2014). Leadership and moral imagination: Beyond the decision-making context. *Rivista Internazionale di Filosofia e Psicologia*, 5(3), 343–353. doi:10.4453/rifp.2014.0027
- Pritchard, M. S. (1998). Professional responsibility: Focusing on the exemplary. *Science and Engineering Ethics*, 4(2), 215–233.
- Pritchard, M. S. (2001). Responsible engineering: The importance of character and imagination. *Science and Engineering Ethics*, 7(3), 391–402.
- Pritchard, M. S. (2006). *Professional Integrity Thinking Ethically*. Lawrence, Kansas: University Press of Kansas.
- Rambo, L. R. (1995). *Understanding Religious Conversion*. New Haven: Yale University Press.
- Reynolds, S. J. (2008). Moral attentiveness: Who pays attention to the moral aspects of life? *Journal of Applied Psychology*, 93(5), 1027–1041.
- Rothman, A. J., Sheeran, P., & Wood, W. (2009). Reflective and automatic processes in the initiation and maintenance of dietary change. *Annals of Behavioral Medicine*, 38, S4–S17. doi:10.1007/s12160-009-9118-3
- Ryan, R. M., & Deci, E. L. (2003). On assimilating identities to the self: A self-determination theory perspective on internalization and integrity within cultures. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of Self and Identity* (pp. 253–274). New York: Guilford Press.
- 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(1), 33–46.
- Schön, D., A. (1984). *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives Emotions, and Behavior: The Better Angels of Our Nature* (pp. 221–241). Washington, DC: American Psychological Association.
- Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and consistency from intra- and inter-personal interactions. *Personality and Social Psychology Review*, 6(4), 316–325.
- Skeem, J. L., Polaschek, D. L. L., Patrick, C. J. et al. (2011). Psychopathic personality: Bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest*, 12(3), 95–162. doi:10.1177/1529100611426706
- Smith, E. R., & Mackie, D. M. (2008). Intergroup emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotion* (pp. 428–439). London: Psychology Press.

- Smith, I. H., Soderberg, A. T., Netchaeva, E. et al. (2023). An examination of mind perception and moral reasoning in ethical decision-making: A mixed-methods approach. *Journal of Business Ethics*, doi:10.1007/s10551-021-05022-9
- Snyder, M., & Clary, E. G. (2000). The functional approach to volunteerism. In G. R. Maio & J. M. Olson (Eds.), *Why We Evaluate: Functions of Attitudes* (pp. 365–393). London: Psychology Press.
- Snyder, M., & Omoto, A. M. (2008). Volunteerism: Social issues perspectives and social policy implications. *Social Issues and Policy Review*, 2(1), 1–36.
- Thalhammer, K., O'Loughlin, P. L., Glazer, M. P. et al. (2007). *Courageous Resistance: The Power of Ordinary People*. Basingstoke: Palgrave Macmillan.
- Van Gordon, W., Sapthiang, S., & Shonin, E. (2021). Contemplative psychology: History, key assumptions, and future directions. *Perspectives on Psychological Science*, 17(1), 99–107. doi:10.1177/1745691620984479
- Vohs, K. D., & Baumeister, R. F. (Eds.). (2016). *Handbook of Self-Regulation: Research Theory, and Application*. New York: Guilford Press.
- Vohs, K. D., Schmeichel, B. J., Lohmann, S. et al. (2021). A multisite preregistered paradigmatic test of the ego-depletion effect. *Psychological Science*, 32(10), 1566–1581.
- Wagner, U., Galli, L., Schott, B. H. et al. (2015). Beautiful friendship: Social sharing of emotions improves subjective feelings and activates the neural reward circuitry. *Social Cognitive Affective Neuroscience*, 10(6), 801–808. doi:10.1093/scan/nsu121
- Wainryb, C., & Brehl, B. A. (2006). I thought she knew that would hurt my feelings: Developing psychological knowledge and moral thinking. In R. Kail (Ed.), *Advances in Child Development and Behavior* (pp. 131–171). Amsterdam: Elsevier.
- Walker, L. J., & Frimer, J. A. (2009). Moral personality exemplified. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 232–255). New York: Cambridge University Press.
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, 86(4), 629–647. doi:10.1037/0022-3514.86.4.629
- Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2013). The integrative process: New research and future directions. *Current Directions in Psychological Science*, 22(1), 69–74. doi:10.1177/0963721412468001
- Weir, K. (2012). *What you need to know about willpower: The psychological science of self-control*. Retrieved from <https://www.apa.org/topics/personality/willpower>
- Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, 15(6), 728–755. doi:10.1080/17405629.2018.1435413
- Williams, B. (1981). *Moral Luck*. New York: Cambridge University Press.
- Winter, M. J. (2011). Does moral virtue require knowledge? A response to Julia Driver. *Ethical Theory and Moral Practice*, 15(4), 533–546. doi:10.1007/s10677-011-9310-y
- Wood, W., Mazar, A., & Neal, D. T. (2021). Habits and goals in human behavior: Separate but interacting systems. *Perspectives on Psychological Science*, 17(2), 590–605. doi:10.1177/1745691621994226
- Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, 21(3), 253–277. doi:10.1177/1088868316652279

Part III

Processes

Final Proofs

Final Proofs

Moral Reason

In every system of morality which I have hitherto met with, I have always remarked that the author proceeds for some time in the ordinary way of reasoning ... when of a sudden I am surprised to find that instead of the usual copulations of propositions *is* and *is not*, I meet with no proposition that is not connected with an *ought* or an *ought not*... . [It] seems altogether inconceivable how this new relation [of *ought*] can be a deduction from others [of *is*] which are entirely different from it.

(David Hume (1739/1985), *Treatise on Human Nature*, 3.1.1)

The grounding of morality in reasoning is often assumed to be central to and even defining of the phenomenon. In this classic statement on the is–ought relation in moral philosophy, David Hume mentions “every system of morality” involving “the ordinary way of reasoning.” Important theorists have claimed that if it does not involve conscious reasoning about the good, then it does not count as moral (Kohlberg, Levine, & Hewer, 1983; Korsgaard, 2006). Thus, any text that purports to cover morality had better talk about moral reasoning. This is the chapter where we do so.

First, we will ask about the nature of reason, and how various models of reason limit the questions we can ask of moral reason. We will then explore in depth the most widely recognized normative model of moral cognition, originating in the research program of Lawrence Kohlberg, who himself claimed his system allowed him to move from the descriptive *is* to the normative *ought*, a move the philosopher David Hume (quoted earlier) thinks is suspicious.¹ While looking at critiques of Kohlberg’s model, we will run across the problem of relativism in morality, a recurring theme in the book. We will then look at more recent research in moral cognition that has arisen

¹See Kohlberg (1971) for this claim, and Moore (1903) for the classic statement why it is a fallacy, specifically, the *naturalistic fallacy*. Though in places cited as received philosophical wisdom, properly speaking it is not a logical fallacy (Frankena, 1939) and the distinction has been called by some simply a “muddle” (Searle, 1964). It is at the least a very interesting muddle. Sayre-McCord (2015) provides a recent overview of the various positions on this issue but seems to think that the burden of proof is on those who want to insist that the naturalistic fallacy is actually a fallacy.

after the demise of the Kohlbergian enterprise. These models open up a world of implicit, nonconscious cognition and multiple possible parallel cognitive and emotional processes associated with moral action. Finally, we will question the role of reason in supporting moral action by recognizing its dark side. We will conclude that current models make possible an understanding of moral evaluation as a complex balance of rational and emotional evaluation of moral action that is usually embedded in specific contexts. This social-emotional reflective equilibrium is required because a balance based in reason alone is still subject to significant bias.

7.1 What Kind of Reason?

What is the role of reason in moral action? This relatively straightforward question has become more contentious in psychology in the last decades. Kohlberg's developmental models of morality (Kohlberg, 1964, 1983) have presumed a basis in formal logic and the ability to generalize it across situations and people (Gibbs, 1977; Gibbs et al., 2007). But other models of rationality parse the world slightly differently. Formal decision theory models of rationality look to maximize our "goal satisfaction" measured by some subjective utility (usefulness) that we perceive in possible outcomes (Payne & Bettman, 2004). We are then rational to the extent that we choose the outcomes that maximize the utility associated with the goal – see Fischhoff and Downs (1997) for an example of this approach and Weber and Johnson (2009) for a critique. As a third option, expertise models of rationality match performance against that of selected experts in the area (Dreyfus & Dreyfus, 2004; Schön, 1984).

In their respective research literature, each of these three approaches usually proceeds without reference to the other approaches, and limits its investigations to those phenomena most congenial to its particular understanding of reason (Evans & Elqayam, 2011). Kohlbergian approaches focus on logical operations in the application of principles (Rest et al., 1999). Decision theory approaches concentrate on problems amenable to probabilistic forecasting (Aktipis & Kurzban, 2004; Sunstein, 2005; Yamagishi et al., 2014). Expertise models focus on skilled performance (Huff, 2014; Narvaez & Bock, 2014; Narvaez & Lapsley, 2005). In the realm of moral psychology this means theorists using one of these approaches limit their understanding of morality to only rational deliberation about morality instead of other possible influences, such as personality or emotion. They then further limit their understanding of morality to rationality considered as either logical operations, utility, or expertise and the particular kinds of moral problem most congenial to that approach.

One tension all the models share is that between prescribing the correct *beliefs* and prescribing correct *actions*. Theoretical reasoning is correct when it leads to correct beliefs (*x* is good, better, best), while practical reasoning is correct when it leads to action that helps to achieve our (im)moral goals (Over, 2004). Most work in judgment and decision-making presupposes practical reason, or "instrumentalist" reason (Over, 2004). Thus: "Reason is wholly instrumental. It cannot tell us where to go; at best it can tell us how to get there. It is a gun for hire that can be employed in the service of any goals we have, good or bad" Simon (1983, pp. 7–8). This "gun for hire" problem will become crucial later as we look at how reason supports immorality.

The instrumental nature of rationality does, however, also allow us to open up the field to see more parts of an action as rational. The cognitive psychologist Over (2004) uses the analogy of passing a car to explain this. When overtaking and passing a car, one could narrate all the preconscious and conscious rational judgments one makes in order to do so successfully. These processes are what are called preconscious; we often do not pay attention to them, but we can monitor them. One also does many other things entirely nonconsciously, such as judging the distance between two cars. It is important to get this judgment “good enough” to support the successful achievement of the goal, but it is usually done below the level of consciousness. One does it with “dedicated cognitive processes” (Over, 2004). The same can be said for an action in the moral realm: being a good partner involves much conscious but also much preconscious and nonconscious rational judgment.² Thus, thinking of reason as *instrumentally* connected to action makes a much broader variety of rationality, judgment, and skill immediately relevant to moral psychology. Conscious moral deliberation becomes just the tip of the iceberg.

Another tension shared by models of rationality is that between normative *prescription* of correct reasoning and simple *description* of how people reason. Many models of reason – and most published research on reason (see Elqayam & Evans, 2011) – presume a *correct* way to reason and this benchmark allows researchers to document ways people depart from the model, thus also achieving the goal of description. But this approach also implicitly restricts its reach to those places where the “correct” answer can be calculated, and thus ignores many domains of life in which people do important reasoning (Evans & Elqayam, 2011). In response, many researchers are now adopting purely descriptive approaches to rationality (for recent attempts in moral psychology, see Killen & Dahl, 2021; Malle, 2021). One lesson for moral psychology is that a turn toward the descriptive may help us get a better grasp of the variety of ways people are moral and the broad range of things that influence moral action.³ But before we move on to these approaches, we will first look at the most widely recognized normative model of moral cognition, one based in conscious logical reasoning about moral principle. Understanding the complexities of Lawrence Kohlberg’s moral psychology and its eventual fragmentation will be crucial to understanding the current state of the field.

7.2 Kohlberg, Relativism, and Moral Reason

If we were required to select the single researcher who had most influenced the development of moral psychology in the last seventy years, the selection would be easy: it is still Lawrence Kohlberg. From his dissertation work in 1958 on moral stages in children, through his long career, and up to his untimely demise in 1987 (Walsh, 2001) Kohlberg founded moral psychology as a sub-field of developmental psychology and guided its research paradigm for three decades. Now, more than three decades after his death, the field is still influenced by his conceptualizations, accumulated research output, and idealistic commitment to moral education (Frimer & Walker, 2008;

²One may consciously plan how to support the other’s plans (conscious) and be listening in conversation for those plans (preconscious), while taking turns in sharing the floor of the conversation (nonconscious). All are relevant to the goal of supporting the other.

³See our argument for this turn in the Introduction.

Thoma, 2014). The cottage industry of research that has flowered in his wake has made significant revisions, and some outright refutations, to his claims. And it has expanded its scope of inquiry far beyond the relatively narrow category of “moral judgment” that his approach prescribed. But even when he was wrong, Kohlberg was usually wrong in interesting ways.

Most presentations of Kohlberg’s system present it as a static six-stage conceptualization (see Box 7.1). On the contrary, it changed constantly in response to Kohlberg’s own data and to challenges from others. For instance, in later longitudinal studies of his original dissertation participants, some of the people *regressed* in his stage model, moving from stages 5 and 6 in high school (the highest) to stages 3 and 4 in college. Kohlberg and Kramer (1969, p. 109) called this “sophomore retrogression.” If one believes stages should not include regression, then this was a fault of the instrument, and the scoring system was revised to reinstate invariant forward progression (Gibbs et al., 2007). This led to a further difficulty, in that the modified scoring system then produced extreme scarcity in attainment of higher stages in the model. Stage 6 ended up being omitted from the final version of the scoring manual (Colby et al., 1983; Gibbs et al., 2007).⁴ Thus, the system was constantly evolving over the years Kohlberg was alive.

This short review of the most well-known version of cognition-based moral judgment will focus on five aspects of Kohlberg’s approach: (1) its challenge to moral relativism; (2) its requirement of explicit judgment; (3) its grounding in cognitive developmental stages; (4) the use of responses to dilemmas as a primary assessment method; and (5) its claims to cultural universality. Each of these still resonate in the field today. We will go on to look at modifications and critiques of the Kohlberg system of moral development.

7.2.1 Kohlberg’s System for Challenging Moral Relativism

This chapter began with the famous comment from the philosopher David Hume that one cannot bridge the gap from descriptive, empirical “is” statements and prescriptive, normative “ought” statements. In an article titled “From is to Ought” Kohlberg (1971) explains his focus on the implications that empirical moral psychology has for philosophy. He claims he wants to go from is to ought and back again in his research primarily because he is worried about how psychological moral relativism might produce manipulative moral education (Kohlberg, 1971, p. 225). In the first pages of his dissertation, he signals his intent to challenge the reigning psychological approaches to moral development (Kohlberg, 1958, pp. 1–3). He views them as relativistic, reducing moral judgment to simple cultural socialization. And this socialization approach does not, in his eyes, recognize the special claim that judgments in the moral realm have – that they are about what is good and right, and that one can transcend and rationally critique a culture’s judgments about the good.

7.2.1.1 Against Moral Relativism One central concern that Kohlberg had with the cultural relativist position was that it gave license to the moral educator to simply manipulate students into accepting the dominant culture’s morality – to socialize them

⁴The collapse of evidence for stage 6 left stranded the even more speculative possibility of a stage 7 in which moral principles were grounded in meta-ethical or religious commitments (Kohlberg & Power, 1981; Walker & Frimer, 2008).

Box 7.1 Kohlberg's Stages

This presentation of the stages is taken from Colby et al. (1983, pp. 3–4). The three major divisions are influenced by Dewey and Tufts and other theorists (Kohlberg, 1971, p. 183).

Preconventional level

Stage 1: Heteronomous morality: Right is judged in concrete, physical terms. Motivation is avoiding punishment. Social perspective is egocentric.

Stage 2: Individualism: Right is judged as acting in one's own interest and being fair. Motivation is serving own interests and recognizing other's interests. Social perspective is concrete and individualistic.

Conventional level

Stage 3: Conformity in mutual relationships: Right is conformity to role expectations and proper motives. Motivation is being a good person in your own and other's eyes and maintaining rules. Social perspective is individual responsibility in mutual relationships.

Stage 4: Accepted duties, societal contribution: Right is conformity to duty constructed within a moral system. Motivation is conscience-based obligation and maintaining the moral system. Social perspective is aware of societal and system level in addition to relationships.

Postconventional level

Stage 5: Postconventional/principled: Right is based on nonrelative values such as life and liberty; these trump rules in social systems. Motivation is obligation to social and interpersonal contract. Social perspective is aware of values that are prior to contracts.

Stage 6: Universal ethical principles: Right is based on universal principles of justice; laws and contracts are right when they rest on these principles. Motivation is personally appropriated universal principles. Social perspective is the foundation of social interaction in moral principle.

The two arrows represent the *parallelism* of the empirically measured developmental pattern (right) with the philosophically normative adequacy (left) of that reasoning (Levine, Kohlberg, & Hewer, 1985, p. 95).

Normative: Higher Levels of Adequacy in Justice Reasoning

Empirical: Measured Transformations of Justice Reasoning

without regard to facilitating critical thinking skills that might generate a critique of the society (Kohlberg, 1971, p. 225). He attempts to ground his response to the cultural relativism dominant in psychology by constructing his system on two kinds of explicitly normative approaches: a *rational* normative approach and a *morally* normative one. It is *rationaly normative* in the sense he borrows from the philosopher John Dewey

(Dewey & Tufts, 1908): action should be based on conscious weighing of reasons based in rationality. Thus, Kohlberg was trying to construct a “theory of rational moral judgment like that now present in economics” (Kohlberg, 1971, p. 225).⁵ It is *morally normative* in that as one progresses in the stages of moral development one incorporates more of this kind of independent, rational, moral reasoning. This progression itself is borrowed from Dewey’s systematization of societal ethical progression, which runs from early group life, to group morality, to personal, principled morality (Dewey & Tufts, 1908, p. 17). They present this in the anthropological categories current at the time as a cultural progression, developing toward a more rational basis for moral judgment in a culture.⁶

But in what way does Kohlberg actually go from “is to ought?” His claim of doing so is not based on a logical derivation. This would be the naturalistic fallacy.⁷ It is instead based on what he described as a “parallelism” between empirical psychology and normative philosophy: that the normative and the empirical will be shown to have similar structure. He claims, for instance, that the empirical developmental sequence he outlines will also roughly correspond to judgments of the ethical adequacy of the reasons given in those stages. In short, the claim is that the higher descriptive stages of the model are philosophically more adequate.

Kohlberg’s project is an attempt to find an empirical argument against relativism in order to limit coercion in moral education. If increasing sequences of human development correspond to increases in moral/ethical judgment, and if this sequence is found cross-culturally, then the philosophical adequacy of moral judgment can itself be said to be both rational and cross-culturally invariant. This parallelism will make cultural relativism difficult to defend. In turn, this will make the idea that morality is simply socialization seem inadequate, thereby putting limits on arbitrary attempts at moral socialization. One can stand up to powerful socializers in the name of moral principle (Killen & Dahl, 2021).

7.2.1.2 *Explicit Judgment* In his attempt to move away from the socialization of morality approach that was broadly accepted in his day, Kohlberg turned away from concepts like conformity and guilt and instead focused on explicit moral judgment, putting the weight of judgment back on the individual: “An action ... is neither good nor bad unless it has been preceded by a judgment of right or wrong” (Kohlberg, 1958, p. 5). In later work (Kohlberg et al., 1983) he would call this the assumption of phenomenism. Any actions that are not consciously chosen for moral reasons are not in the domain of study. This assumption has until recently been widely shared in moral psychology (Narvaez, 2010). This approach treats nonconscious influence, personality, emotion, and even skilled automatized action as at best irrelevant and at worst a biasing influence (Walker & Hennig, 2004). The current blooming of research

⁵See also p. 52 in his 1958 dissertation.

⁶These cultural assumptions have been strongly critiqued in the current anthropological literature (Keane, 2015; Shweder, Mahapatra, & Miller, 1987).

⁷See n. 1.

in moral psychology has come about in part because it has freed itself from this narrow definition. But the narrow focus on conscious reason was central to Kohlberg's project of overcoming the relativism inherent in the pure socialization approaches of behaviorist and psychodynamic approaches. This suggests that rejecting Kohlberg's conscious phenomenalism may come with a cost: opening up descriptive psychological approaches to a naïve cultural relativism. For instance, the recent work in moral foundations theory is quite explicit in its interest in cultural differences in morality and embrace of relativism (Graham et al., 2011). Other theorists are skeptical of this turn (Killen & Dahl, 2021; Narvaez, 2010).

7.2.1.3 Cognitive Developmental Stages In constructing his developmental stage system, Kohlberg put two keys from Piaget to good use. First was his construction of stages that were cumulative, invariant, and dependent on the logic of the domain for their sequence. The second followed from this, and was Piaget's identification of social interaction as a primary developmental influence in moral development (Gibbs et al., 2007).

Thus, the stages identified in Box 7.1 do not, in Kohlberg's system, depend on any specially dedicated evolutionary modules (Cosmides & Tooby, 2008) or on anything more than biological development and accumulated experience that support the ability to make the relevant distinctions. Once one has the cognitive capacity to recognize others' perspectives, and enough practice doing so, one can move from pre-conventional to conventional moral reasoning. Thus, one can expect that advancement in moral reasoning stages will co-occur with advancement in abstract reasoning, but also with more extensive experience of peer interaction that requires and provides practice for social perspective taking. Moral judgment and its development are patterned by the kind of cognitive and social creatures we are. One outcome of this emphasis is that, at least at the lower levels, one can expect the stages and their progression to be cross-culturally similar. But to the extent that cultures restrict opportunities for social perspective taking (by, e.g. restricting social interaction), one might see less movement to higher stages – specifically because it is in part this social interaction that drives development.

This commitment to stage progression based on logical properties of the domain allowed Kohlberg to challenge relativism by showing that, across cultures, more philosophically adequate moral reasoning was achieved by a naturally occurring developmental sequence. But its cognitive foundations also required the model to assume that advances in the logical progression could not be retreated from. Once one acquires the capacity and learns the logical transformation of perspective taking, regression away from this ability makes little sense, unless it is cognitive capacity that is being lost.⁸

⁸At least, it makes little sense in a logical system, and threatens to undermine the anti-relativist goal of the system. It might well make sense in a human system. But to recognize this would lose the parallelism that is central to the project.

7.2.1.4 Dilemma-based Assessment Kohlberg's assessment model was, as fits his system, based in the conscious giving of reasons in a conversation. But it was a particular kind of conversation he had most in mind. As he initially said in his dissertation:

it is clear that our conception of morality implies or refers to rather complex thoughts and feelings about action in genuine conflict or crisis situations. Accordingly, we decided to use extensively probed open-ended individual interviews about hypothetical conflict situations which posed genuine dilemmas to educated adults (Kohlberg, 1958, pp. 75–76)

Thus, instead of using Piaget's method of asking for comment on pairs of stories that are modified in particular ways (e.g. a child broke fifteen cups vs. just one cup), Kohlberg selected stories where values were conflicted and one had to choose a course of action. Which action was chosen was not central to the coding. Kohlberg rather concentrated on the *reasons* one gave for the choice and how one reasoned. This assessment method was cumbersome, time-consuming, and required highly trained interviewers and coders. Others have adapted the dilemma technique to questionnaire forms that are machine-scored (Rest et al., 1999), and this achievement is in part responsible for the widespread use of Kohlberg's approach as a method in ethics assessment.

There is another legacy to Kohlberg's measurement approach: the giving of rational explanations in the context of value conflicts and choice dilemmas. This looks most like what philosophers do in constructing cases to make their arguments in ethics and meta-ethics. But because it always places the thinker in a difficult moral choice, it narrows the domain of ethics and morality to what Pincoffs (1971, p. 571) has called "quandary ethics," and it results in "an indefensibly narrow conception of the subject of ethics." This leaves us with an even narrower focus in morality: not only conscious reasoning about morality but also conscious reasoning about a particular aspect of morality – those situations where values are in conflict and the only possible options play those values off against each other. This approach ignores moral action that is done without conscious intent, but it excludes even the moral cognition and action of exemplars who spend most of their conscious (and perhaps unconscious) energy pursuing a moral goal to which they long ago dedicated themselves.

7.2.1.5 Universality Claims Kohlberg's system is founded in the development of basic human reasoning capacity in the context of social perspective taking and reason giving. Because it has this structure, it makes a controversial claim for the cultural universality of the stages of moral reasoning. All people in all cultures should move from one stage to the next as they develop in moral reasoning. And this development in moral reasoning should be tied to practice in social perspective taking. This leads to two empirical predictions: (1) that, at least in early stages there should be parallel development across cultures; and (2) that this development, particularly for higher stages, should depend on social experience that requires perspective taking and reasoning about it. An implication of this approach is that:

moral judgment maturity should correlate with age, education, higher socioeconomic status, urban (versus rural) settings, and community participation, insofar as these variables index the affordance and accumulation of diverse social experiences and perspective-taking opportunities through social participation (Gibbs et al., 2007, p. 446)

The short answer to this question is that, at least for the progression in stages 1–3, one does find the correlations one would expect across a wide variety of cultures: two large reviews of the evidence in studies across forty-two countries come to this conclusion (Gibbs et al., 2007). Progression from stages 1 to 2 occurred in most cultures by late childhood, with the exceptions having to do with experiential variables (e.g. education, rural setting) that would reduce exposure to social perspective taking. By late adolescence, stage 3 becomes the predominant reasoning level across cultures, and stage 4 begins to predominate among adults across cultures, with the exception of rural and less well-educated samples.

But at the higher levels, particularly the transition to post-conventional thinking, the stage progression begins to fall apart. It seems most descriptive of the development of individuals, across cultures, with education at the university level, perhaps because this facilitates social perspective taking (Gibbs et al., 2007). Does this mean that individuals in rural cultures are necessarily less morally developed in their reasoning? Or that the theory is necessarily biased against those without academic training?

Some theorists have suggested that the emphasis on justice reasoning in Kohlberg's post-conventional stages misses the diversity of ways one might become post-conventional in morality. One can critique and transcend conventional moral standards, for instance, by existential reflection (Gibbs, 1977), or by using the resources already available in the culture to propose alternative ways of thinking (Gibbs et al., 2007). Asking "Why do bad things happen to good people?" or "Why be moral?" can be the beginning of stepping outside of conventional morality and facing moral questions from an independent (and thus post-conventional), existential perspective. But one can also use religious principle to critique conventional morality. One of the central characteristics of religion is its critique of conventional morality (Bellah, 2011). Gregg (2005, chapter 3) reviews how the values of Islam help to counter and limit the honor/shame system in Middle Eastern countries, and even help to "set the developmental task of identity formation ... to integrate pursuits of honor and piety into a whole, or at least a balanced life" (Gregg, 2005, p. 167).⁹

7.2.2 Critiques of Kohlberg's System

In part because of his controversial claims, and in part because the system was for many years the reigning paradigm for thinking about morality, Kohlberg's system has collected many detractors, critiques, and attempts at modification. We have already seen evidence of two, the limitations of conscious rationality and the complexity of cross-cultural application. Here we will review a few others that have received attention.

7.2.2.1 Gender and the Concept of Care One of the most widely known controversies in moral psychology of the previous century was that initiated by Carol Gilligan in her

⁹Kohlberg would score basic religious commitments as a part of conventional morality (Richards & Davison, 1992) but might score more sophisticated religious reflection as post-conventional (Gibbs, 1977).

(1977) book *In a Different Voice*. Its strength for moral psychology was to point out an alternative developmental path for moral judgment and action that focused on human relations and an ethic of care. Gilligan claimed that women were more guided by care in networks of relationship while men were guided by abstract concerns of justice and thus that Kohlberg's system underestimates the moral development of women. These claims did not hold up under further investigation (Gibbs et al., 2007; Walker, 1984). If anything, there is a slight developmental advantage for women, and both men and women use care reasoning about equally.

7.2.2.2 Pre-Conventional Morality and Complexity The early stages of Kohlberg's system have also suffered from fragmentation. Young children are often genuinely other-oriented, though they have trouble connecting this with reasoning about others' experience and their own motivations (Malti & Ongley, 2014). Even toddlers (Tomasello & Vaish, 2013) and infants (Van de Vondervoort & Hamlin, 2018; Wynn & Bloom, 2014) understand something about helping, cooperation, other's feelings, etc. Much of the development is tracked by development of "theory of mind," that is, the child's understanding of the relations between mental states, emotions, and behavior in both the self and other (Lagattuta & Weller, 2014; Malti & Ongley, 2014; Wellman, 2018).

7.2.2.3 Post-Conventional Morality and Cultural Pluralism The anthropologists Shweder and Menon (2014) agree with Kohlberg that a simplistic cultural relativism poses significant problems for a good description of the special status of morality across cultures but critique Kohlberg's approach to moral objectivism. Their moral objectivism is instead a descriptive moral universal: all cultures seem to assume an "objective moral charter." Across all cultures, moral judgments are seen as making claims about an objective moral reality that other reasonable minds must agree with. The relativist ideal means in the end that "there is no meaningful distinction to be drawn between morality and power" whether that power is exercised by individuals, governments, or cultures (Shweder & Menon, 2014, pp. 14–15). Thus, they argue for a pluralist vision, one that holds in tension cultural diversity of moral judgment, respect for the moral judgment of others, and a commitment to understanding the "objective moral charter." Of course power is involved in social change attempts, but it is often not simply about power but about a vision of the good (Killen & Dahl, 2021).

7.2.3 What Remains of Kohlberg

A systematic response to these difficulties in the Kohlberg project has been mounted by James Rest, Darcia Narvaez, Muriel Bebeau, and Steven Thoma in what they call a "neo-Kohlbergian" model of morality (Rest et al., 1999). Their approach is consciously eclectic, borrowing from multiple research traditions within psychology (Narvaez, 2005, 2010; Narvaez & Bock, 2002, 2014; Thoma, Derryberry, & Crowson, 2013) to construct a recognizably Kohlbergian approach that is much more flexible and congenial to the current burgeoning of research in moral reasoning presented in the rest of this chapter.

One thing not modified in the neo-Kohlbergians' various accounts is the commitment to a Kohlberg-style moral objectivism rooted in a developmental worldview: it

is difficult to talk about moral *development* if one thinks any moral approach is as adequate as the next. Development beyond childhood requires a conventional recognition of mutual dependence and social rules that limit or modify one's desires. Further development beyond this conventional stage requires the ability to construct a representation of that system and to critique it. This cultural critique is made possible by natural variations in value systems in the moral ecology that can provide cognitive and emotional resources for both critique of conventional morality and moral persuasion to influence the conventional moral ecology. This critique and persuasion can use such cultural resources as religion, social influence from groups, shared cultural commitments (e.g. agreed upon rights), and attractive practices from other cultures (e.g. recognition of various rights) to critique the culture itself and to work for social change. At the individual level, it may also be these resources that make possible a switch to a post-conventional moral commitment. There is still much work to be done to document how this occurs (Keane, 2015) and whether this critique depends psychologically on an objectivist conception of moral judgment (Killen & Dahl, 2021).

Without this ability to engage in cultural critique "there might still be an Atlantic slave trade" (Narvaez, 2010, p. 168). Narvaez and colleagues see this sort of moral advance as the result of "more agile perspective taking, an ability to appeal to ideals that are shareable and not exclusive" (Narvaez, 2010, p. 167). The context of this quote suggests that this "post-conventional thinking" should be shareable across cultures and thus provide a limited and negotiable kind of cultural universal in the domain of morality. The neo-Kohlbergian approaches see cognitive, schematic processing as an important influence in moral judgment and action. However, they also see intuitive, implicit processing and emotion as equally important partners. It is to these intuitive aspects of cognition that we now turn.

7.3 Implicit Cognition and Two-Process Models

That we are often "of two minds" about things has been recognized by both ancient and modern authors. Augustine of Hippo's *Confessions* (Augustine, 397/2007, Book 8, chapter 9) offered an early exploration of multiple wills. The Scottish empiricist Adam Smith (1759/2009) proposed that it was the interaction of our passions and an imagined "impartial spectator" that together shaped our choices. Schelling (1800/1993) revived the concept of the unconscious as a central part of knowing and acting (McGrath, 2012). Kierkegaard explored the dynamics of multiple aspects of the self (Kierkegaard, 1849/2004). And, of course, Freud systematized the unconscious (Breuer & Freud, 1895/2000) as a dynamic, emotionally rich driver of thought and action. All these theorists were pointing to various aspects of unconscious processing and its relation to conscious, controlled processing.

More recently, the unconscious has been resurrected by cognitive psychologists as *nonconscious processing*, at least initially stripped of its association with creative impulses and emotional content. Fodor (1983) introduced one of the first such systems, though the term "dual process" had been used earlier in reference to a contrast between conscious and nonconscious influences on behavior (Wason & Evans, 1975). Since these early beginnings, work on dual process theories of cognition

could fill volumes (e.g. Kahneman, 2011; Nisbett & Ross, 1980). Dual process theories are widely divergent in their assumptions (see Evans & Stanovich, 2013 for a review) and have some detractors (Kruglanski & Thompson, 1999; Osman, 2004). But they have proven quite influential in moral psychology. After looking for some structure in what we might mean by dual processes, we will review four significant theoretical approaches in moral psychology that have been influenced by dual process theories but give different answers to what the content of those processes might be.

7.3.1 What Are Two-Process Models?

It is important to begin this section with the plural. There is not one agreed upon two-process approach to understanding cognition. Theorists disagree on a range of issues, including whether there are two or even more systems, whether the difference is one of type (with a clear distinction) or of mode (with a range of intermediate positions), the extent to which they can be consciously controlled, and a host of other disagreements (Evans & Stanovich, 2013). Any generalization at this early stage of theorizing should be done with caution.

Nevertheless, if we assume for the moment that there are only two types of processing, and that they are distinct, what is the distinction? It is that some of the cognitive processing we do requires cognitive effort – it takes concentration and working memory. But at other times, and for other kinds of stimuli, processing is relatively effortless and requires little in the way of working memory. This latter kind of process has been called “system 1” processing, in partial reference to the thought that it evolved earlier than the more effortful, working-memory-intensive processing, which is called “system 2.” Table 7.1 presents a schematic of the two systems of processing. The items listed under *defining features* are, according to Evans and Stanovich (2013), the characteristics of the two types that are central and always present, while the *typical correlates* section lists characteristics that are often associated with the type, but for which there are exceptions.¹⁰ The bottom section of the table lists characteristics of the two processes that are often cited when the processes are proposed as different systems that evolved somewhat separately.

Like Kohlberg (1958), when we think of cognition we normally think of conscious, deliberate, guided processing, or type 2 processing. We usually have conscious access to both the processes and their outcomes. That is, we are aware of and deliberately guide our selection of the things we consider, how and how long we consider them, and the conclusions we reach. This deliberation takes effort, can be abstract and hypothetical (e.g. What if it were not my country’s flag?) and is associated with intelligence. Type 1 processes are rapid, more hidden from our awareness, usually proceed without our conscious guidance, and we only have access to their seemingly self-evident conclusions. For this reason, type 1 processes are often called “intuitive.”

¹⁰This slipperiness of the definitions of the two systems is not new. As early as 1989, Bargh argued that awareness, attention, intention, and control are a loosely tethered set of attributes of automatic processes. Thus, the idea that type 1 processes occur in great variety has at least a moderately long history.

Table 7.1 Characteristic properties of dual-process cognition.

Type 1 process (intuitive)	Type 2 process (reflective)
Defining features:	
Does not require working memory	Requires working memory
Autonomous	Cognitive decoupling; mental simulation
Typical (but not necessary) correlates:	
Fast	Slow
High capacity	Capacity limited
Parallel	Serial
Nonconscious	Conscious
Biased responses	Normative responses
Contextualized	Abstract
Automatic	Controlled
Associative	Rule-based
Experience-based decision-making	Consequential decision-making
Independent of cognitive ability	Correlated with cognitive ability
When considered as evolved systems:	
Evolved early	Evolved late
Shared with other animals	Distinctively human
Implicit knowledge	Explicit knowledge
Basic emotions	Complex emotions

Source: Adapted from Evans and Stanovich (2013) and De Neys (2021).

A good example of how the two processes interact in the moral domain is *moral dumbfounding* (Haidt, Björklund, & Murphy, 2000; Haidt, Koller, & Dias, 1993). This occurs when we have an immediate, obvious, moral reaction to some occurrence (e.g. someone in private cleaning the toilet with the flag of their country).¹¹ In this reaction, one can see the two processes take somewhat separate paths. The immediacy of the initial moral reaction is a clue that we are dealing with a type 1 process. But when asked to give *reasons* for the reaction, people often spin their mental wheels, unable to overcome the parts of the scenario that suggest it should not be condemned (e.g. it is a private act, based in practical frugality, no one knows of it, it hurts no one). To answer “it’s just dishonorable” feels to many people like it is not a real, logic-based reason.¹² Thus, those considering this scenario are often left dumbfounded; they have a clear, compelling moral reaction (type 1) but cannot generate a (type 2) reason for it. The two processes come apart and it appears that the type 2

¹¹It should be noted that Western, liberal college students are the least likely to have moral objections to the flag scenario (Haidt et al., 1993). But for many, in many cultures, it produces an immediate negative, condemning reaction.

¹²Note that this is a descriptive account of the experience of subjects who have participated in these experiments (and of our students’ self-report in classes). Under many philosophical accounts and in many cultures, honor can indeed be a moral reason on its own (e.g. Bok, 2014).

conscious reason system is left in the dishonorable position of futilely trying to give reasons for a moral judgment that has already been made.

7.3.2 Complexities of Two-Process Approaches

Now that we have the basic outline, we can begin to explore some of the complexities. Most of those complexities have to do with type 1 processing and the interactions between the two types of processes.

Some theorists suggest that in the moral domain, type 2 reasoning is not the “real” cause of our moral judgments but only a servant of the intuitive, more emotional type 1 processes. As Hume put it in the eighteenth century, “Reason is, and ought only to be, the slave of the passions” (Haidt, 2001; Hume, 1739/1985, section 2.3.3). Unfortunately, much of the debate about whether type 1 or 2 reasoning is the primary source for moral judgment is misplaced because it misses the profound interaction between the two processes. In Chapter 6 we can see how practice can transform a conscious process of judgment into an automatic one, and how cognitive disfluency can recall an expert from automatic into controlled conscious processing of a difficult problem. In both childhood and adulthood we can train and influence type 1 processing, partly based in our type 2 decision-making and planning.¹³ Thus, the intuitions of type 1 processing provide grist for the mill of type 2 processing, and the planning and decision-making of type 2 processes help to construct experiences that train and influence type 1 processes. In such an interdependent system, attempting to privilege one system over the other as *the* source of moral cognition for all or most instances is futile.

However, people may well have individual differences in the extent to which they use and develop one system or the other (Evans & Stanovich, 2013). For instance, the work of Tversky and Kahneman (1974) has been a long campaign of showing bias in people’s thinking in a wide variety of cognitive tasks where there is a known (but nonintuitive) right answer. This was some of the original work that motivated theorists to develop two-process theories of cognition. But there is a group of people who consistently do better at these tasks: those who are appropriately motivated and people higher in intelligence. Still, even those high in intelligence use the “biased,” intuitive, system 1 approaches when they provide the correct answer. Thus, some people do privilege one system over the other, at least some of the time. But we cannot easily maintain that one system is more basic or more biased than the other. They interact, are both subject to bias, and can be thought of as balancing each other.¹⁴

7.3.3 Contents of Type 1 Processes

So far, we have only talked about characteristics of type 1 *processes* (e.g. automatic, fast, nonconscious). But this leaves unanswered the question of what the actual *contents* of these processes are. What are people (not) thinking about when engaging in type 1 and type 2 processes in the moral domain?

¹³We cover the processes underlying this training in Chapters 6 and 9.

¹⁴See Section 7.5.3.

We review here four somewhat independent answers to this question.¹⁵ The first makes the claim that strongly content specific modules have evolved that, in effect, *do our moral reasoning for us* in particular domains. The remaining three approaches sometime use this module language but mean something less stringent by it. All the approaches assume that there is some set of principles, values, or schemas that structure both type 1 and 2 ways of thinking about morality. They are organized in advance of experience but not necessarily tightly organized and automatic.

7.3.3.1 Moral Modules as Content Many studies show a regularity in the way people ignore logical structure and reason within specific moral content areas (e.g. cheating, incest; these are called “content effects”). Some theorists claim that content effects show the primacy of strong evolutionary modules for content-specific reasoning (Cosmides & Tooby, 2008; Cushman, Young, & Hauser, 2006).

For instance, Cosmides and Tooby (2008, p. 65) approach the contents of moral reasoning from what they call an “ecological rationality” standpoint, consisting of the broad claim that there is a limited set of “functionally isolated” modules for moral rationality. Each has evolved separately, is associated with a particular domain (cheating, dominance, hazard detection), and operates independently of oversight by conscious type 2 reasoning (Beller, 2010). Each module independently evolved because that set of reasoning procedures was an “evolutionarily stable strategy” that consistently gave survival advantage to the species (Cosmides & Tooby, 2008, p. 70).

Evidence for these structures is provided by the established effect that people are very good at detecting and reasoning about cheating (John was obligated to share, he did not). In logical form, this is a violation of an if-then, or conditional, rule. But while they are particularly good at detecting cheating, people are not good at detecting all violations of if-then rules. This at least establishes that cheating is part of a special domain where people perform if-then violation detection very well, setting it apart from general reasoning ability.

This strong claim that there are independent modules of particular types of moral reasoning has been disputed by both philosophers (Prinz, 2008) and psychologists (Chater & Oaksford, 1996; Pietraszewski & Wertz, 2022). The most cited claim is how difficult it is to establish the evolutionary argument (see Chapter 1). The strong claim for independent, evolutionary modules of moral reasoning remains controversial. But there are more flexible kinds of evolutionary modules that require a less stringent evidence base. We turn now to these.

7.3.3.2 Moral Foundations as Content In a program of research originally focused on moral emotion and culture (Haidt et al., 1993; Rozin et al., 1999), Jonathon Haidt and colleagues have done much to establish intuitive type 1 reasoning as a foundational feature of human moral cognition. The intertwining of emotion and reasoning in this approach is so crucial that we will cover the theory in both this chapter and Chapter 8.

¹⁵There are, of course, more approaches than four. For example, Schein and Gray (2015) insist there is only one dimension, harm, and Janoff-Bulman and colleagues include a dimension to explain why we might have different morals regarding in- vs. out-groups (Janoff-Bulman & Carnes, 2013; Janoff-Bulman & Carnes, 2018; Janoff-Bulman, Sheikh, & Baldacci, 2008; Janoff-Bulman, Sheikh, & Hepp, 2009).

A deep connection of emotion and cognition can be found in moral foundations theory, an approach that is an explicit rejection of the rationally based approach found in Kohlberg. This suggests the distance psychological moral theory has recently traveled.¹⁶ For now, we will concentrate on those aspects of the theory that are typically type 1 rationality.

Haidt and colleagues claim that moral reasoning does not typically consist of conscious, guided, logical reasoning based in principles (Haidt, 2007). It is instead based in nonconscious processes. These processes are structured by our evolutionary history, and they produce an emotion-based conscious awareness that something is wrong or right, but without any access to the processes that produced reaction. Thus, moral judgment is typically and foundationally an intuitive type 1 process. This much seems to be shared with the strong moral module approaches covered earlier. But the moral foundations here are not strictly limited to specific domains like incest or cheating (though they might be best triggered by stimuli from particular domains). They are more general ways of responding to the environment. From a cognitive standpoint, moral foundations can be seen as moderately flexible schemas that might be activated in a range of situations and then produce different moral reactions depending on which foundation is activated.

In the evolution of this approach, one can find four (Haidt & Joseph, 2004), five (Haidt, 2007), or six (Iyer et al., 2012) separate foundations. The most often-cited formulation includes five foundations, each labeled with two words to indicate the range of meaning:

1. *Harm/care*: A concern for the well-being of individuals and whether they are harmed or helped.
2. *Fairness/reciprocity*: A concern for the equal or equitable treatment of people, including issues of rights, distribution of resources and privileges, and liberty.
3. *In-group/loyalty*: A concern for loyalty to and maintenance of a group, including family, organizations, friendships, teams, ethnicities, and countries.
4. *Authority/respect*: A concern for respect and deference to authority or privilege as an obligation in social relations.
5. *Purity/sanctity*: A concern for the sacred, the pure, or the divine that may be as variously expressed as cleanliness or obedience to divine command.

Foundations 1, 2, 4, and 5 were the early list proposed, and a sixth foundation of liberty/oppression has been recently proposed. Haidt and colleagues see this fluctuation as part of an empirical attempt to find a set of schemas that are foundational in that they have an evolutionary heritage and structure moral reactions in a distinct way (Graham et al., 2011). Their pragmatic approach is based on seeing the structure of morality as an empirical question. Though they usually list five foundations, they are open to additional ones as well as to evidence that suggests that two or more of them are in fact one unit. Much of the research to support these foundational schemas as

¹⁶If one looks further back than Kohlberg, one can find that the new two-process approaches share much with learning and psychodynamic approaches, though couched in very different language.

indeed foundational comes from large-scale internet surveys and self-report measures, though some have looked at existing texts of conservative and liberal sermons for patterns of their use (Graham, Haidt, & Nosek, 2009) and investigated the relationship between the intuitions that proceed from the foundations and conscious awareness of their use (Haidt, 2001; Haidt et al., 1993).

The causal status of the foundations is somewhat ambiguous, since much of the research involves direct self-report of reasoning associated with each foundation – though there is evidence that they operate at a nonconscious level (Helzer & Pizarro, 2011; Wheatley & Haidt, 2005). The foundations are, then, associated with both type 1 and 2 cognition and it is unclear which aspect is more important. Nor is it clear that the foundations Haidt and colleagues list as typically liberal (1 and 2) and conservative (3 to 5) are necessarily at odds. Frimer, Tell, and Haidt (2015), for instance, find evidence that for certain issues (e.g. the environment) sacredness is shared among conservatives and liberals. Though there is still much work to do in determining how system 1 and 2 processes interact with schema processing, the moral foundations approach has become a central research program in moral psychology.

7.3.3.3 Human Values as Content Any theory of moral reasoning depends on a theory of value that will identify the moral good (Schroeder, 2021). Thus, another approach to the contents of moral reasoning is the rigorously inductive, cross-cultural mapping of values by Schwartz and colleagues (Bardi & Schwartz, 2003; Roccas et al., 2002; Schwartz, 2006, 2010).¹⁷ They have shown a two-dimensional organization that defines the value dimensions within which individuals from differing cultures operate.¹⁸ Work in over seventy countries using a variety of instruments converges on this solution that represents the similarities in the way values relate to each other across cultures. Figure 7.1 shows the standard circular arrangement of the 10-value solution, with the two dimensions labeled self-enhancement – self-transcendence and conservation – openness to change. At both individual and at cultural levels, support of any one value on the circle tends to correlate with support for adjacent ones, and less support for those on the opposite pole. As with other attempts to map the structure of moral concerns, the world of morality does not always fall into easily discerned categories. Depending upon the instrument, populations, and criteria, Schwartz and colleagues have suggested seven (Schwartz & Bilsky, 1987), ten (Schwartz, 1992), and nineteen (Schwartz et al., 2012) values. There is extensive overlap between the foundations approach of Haidt and colleagues and the value approach of Schwartz and colleagues (Graham et al., 2011; Zapko-Willmes, Schwartz, Richter, & Kandler, 2021), but some domains, such as sacredness, seem unique to the foundations approach while others (e.g. hedonism and more recently, concern for nature) are uniquely represented in the values approach.

The relation of these values at cultural and individual levels is still an active area of inquiry (Borg, Bardi, & Schwartz, 2015; Schwartz, 2016). One anthropological approach is to speak of more or less local morality systems that are particular

¹⁷We also cover this approach in the chapter on Moral Ecology.

¹⁸Work by Geert Hofstede (2001) and by Ronald Inglehart and Christian Welzel (2005) makes similar points about variation in values among cultures.

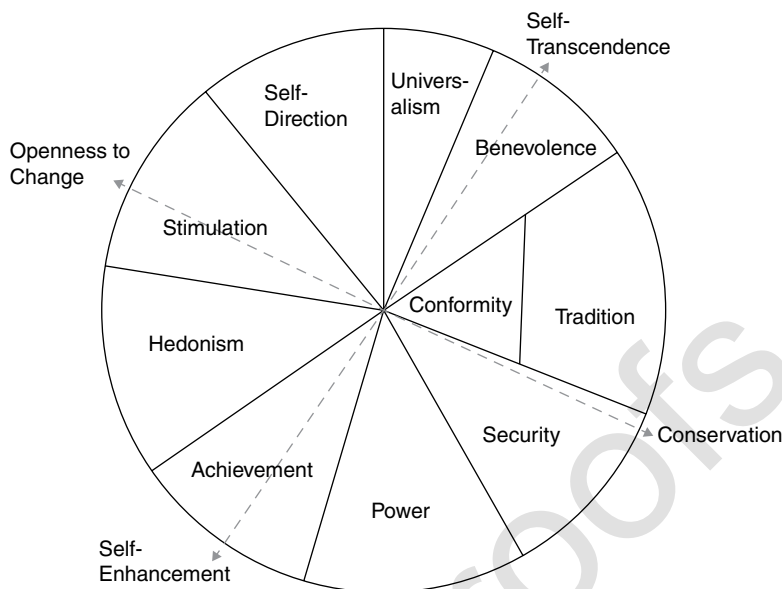


Figure 7.1 Schwartz circle of value regions.

Source: Adapted from Borg, Bardi, and Schwartz (2015).

configurations of the values, with their complex embodiment in the artifacts, ceremonies, documents etc. of a culture, and with the individual negotiating among the multiple morality systems that they encounter (Keane, 2015).

7.3.3.4 Relational Models as Content A final approach to the contents of moral reasoning draws its methods and criteria from anthropology. Fiske and colleagues (A. P. Fiske, 2002; A. P. Fiske & Haslam, 2005; Rai & Fiske, 2011) base their approach in the judgments that are appropriate to certain kinds of social relations. The advantage of this approach is that it makes clear its particular theoretical stance: that moral judgment is rooted in regulating obligations and privileges in social relations, and that different kinds of relations have different roles, obligations, and privileges. The approach is thus rigorously descriptive rather than normative. It describes the roles and the judgments that fit with them. What are the dimensions on which people agree and disagree in particular kinds of relationships? How are the dimensions used? How do the judgments structure social relations (and vice versa)? The approach thus avoids talk of “bias” in moral reasoning. Here, the task is simply to describe how the judgments cluster based on the social relations they support and how they function. The four basic relations emerge from extensive evidence and theory in cross-cultural work in anthropology and consist of:

- **Unity – basic relation:** in-group
Motive: care for and support the integrity of in-groups by eliminating contamination threats and providing aid based in compassion.

- Hierarchy – *basic relation*: hierarchy
Motive: respect rank in social groups. Superiors are entitled to deference but have obligations to the group.
- Equality – *basic relation*: equal footing
Motive: balanced, in-kind reciprocity, equal treatment, input, and opportunity.
- Proportionality – *basic relation*: proportional or economic exchange
Motive: rewards and punishments should be proportionate to merit, benefits to contributions.

These roles can often be mixed in particular relationships. Two friends might communally share some items, trade others, and let one take the lead in an area of expertise.

In addition to describing the social relations underlying morality and the motivating power of morality, Fiske claims to provide a universal structure for the agreements and disagreements among individuals within a culture and between cultures. Cultural disagreement is ubiquitous but has a structure. Any act, from cutting one's hair to genocide, might be justified or condemned in moral terms within a culture at a particular time by being placed in relation to one of these motives. Thus, Fiske makes the claim (also suggested by others, e.g. Graham & Haidt, 2012; Skitka, Hanson, & Wisneski, 2016; Skitka & Morgan, 2014; Skitka & Mullen, 2002) that even what some call evil acts can be seen by their perpetrators as moral obligations, or at least as morally permissible options.

The claim is that cross-cultural understanding and influence is made possible by understanding the moral basis of the claims that individuals in each culture have. Fiske labels his approach a pluralist one, meaning it recognizes more than one value by which one measures moral goodness. But the approach does not seem to have any commitments within it that would keep one from a thorough relativist conclusion – that all morality is reduced to power struggles based on these dimensions. An alternative, though still descriptive, anthropological approach is offered by Shweder (Shweder & Menon, 2014; Shweder et al., 1997), which recognizes and attempts to grapple with an “objective moral charter” that sees moral judgments as different from simple influence attempts based in power.

7.3.4 Critiques of Two-Process Models

We should not end this section without noting that despite their clear success in capturing the market of research in moral cognition, two-process models have also drawn some criticism, much of it trying to maintain a single system (Hassin, 2013; Hassin, Bargh, & Zimmerman, 2009; Kruglanski & Thompson, 1999; Osman, 2004). Evans and Stanovich (2013) provide convincing replies to critiques of two-process models of cognition. One outcome of their defense of multiple processes is that there may indeed be more than two types of processing (taking us back to the evolutionary modules with which we began this section).

Malle (2021) has reshuffled the types of moral judgments to suggest a different path forward. If, as it seems, there are at least four types of moral judgments (evaluations, norm judgments, wrongness judgments, and blame judgments), each with different objects, information, speed, and social function, one might take the characteristics of

each of the two processes (speed, access, automaticity, control, etc.) and do the empirical work to ask which characteristics are associated with which types of moral judgments (Malle, 2021, p. 310). This sidesteps the entire two-process approach and at the same time offers an empirical test of it. This approach might find evidence for more than two processes, with different relations to different types of moral judgment and interactions among them.

De Neys and colleagues (Bago & De Neys, 2019; De Neys, 2012, 2021) offer another approach to fragmenting the two processes. They provide convincing evidence that system 1, the supposedly nonlogical intuitive system does indeed provide (1) intuition based on logical inference and (2) an intuition of potential conflict between that inference and any associative intuition (see also Hassin, 2013; Hassin et al., 2009). This intuition of conflict relates interestingly to the work covered in Chapter 6 on the metacognitive disfluency experts use to signal the need to drop out of automatic skilled behavior and begin problem-solving. System 1 (or whatever name one now gives it) might be providing this essential disfluency intuition in expertise.

From an applied social change perspective, Killen and Dahl (2021) complain that two-process approaches in practice restrict the emotional aspect of intuition to negative affect, not considering compassion, a motivation that is clearly in evidence in many social change movements (and not likely generated by system 2). In addition, the relatively narrow focus on the truth value of the “outcome” of the processes fails to account for the urgency and deep commitment that must go along with any judgment that is to result in a prolonged and committed attempt to produce societal change (Killen & Dahl, 2021, p. 1216). Section 7.4 is relevant to this critique.

Finally, Cesario, Johnson, and Eisthen (2020) suggest that most presentations of two-process systems ignore the consensus in the neuroscience community that there are no simple divisions into automatic and online systems and that characterization of system 1 as an “early evolutionary” system ignores significant continuity among species and unnecessarily separates many different cooperative systems.

Because of this tremendous flux in data and theory, the reader would be ill-advised to conclude that there are only two cognitive processes and that those are mostly differentiated on whether they are open to consciousness or not. More than likely, there are multiple processes that share similarities and differences that influence a variety of types of moral cognition. And one probably find things influencing moral cognition in almost all of them (Dinh & Lord, 2013).

7.4 Naturalistic Moral Cognition

Another promising approach to moral cognition shares at least one assumption with two-process approaches: that rather than depending on a general moral logic, domain-specific cognition matters. These approaches have variously been called naturalistic decision-making (Lipschitz et al., 2001), real-life decision structuring (Galotti, 2007), grounded cognition (Barsolou et al., 2003), and autobiographical reasoning (De Silveira & Habermas, 2011). Here, we refer to them all as naturalistic decision-making

(Zsombok & Klein, 2014). Research in these overlapping areas shares several features (Evans & Elqayam, 2011):

- It is focused on capabilities that people possess rather than a model of how one ought to think (see Section 7.1).
- It attempts to be descriptive of the actual processes based on those abilities.
- It is not usually done in labs, or with naïve subjects.
- The situations of interest often involve severe time pressure (e.g. military decisions), ill-defined goals (e.g. college selection), and high personal stakes (e.g. firefighting).
- It is shaped by the structure of the environment as much as the abilities of the actor.
- It often involves studying people with some degree of expertise.

Research like this finds that perceiving situations and matching roles or solutions is more important than calculating outcomes or courses of action. For instance, chess experts match possible solutions to situations (Newell & Simon, 1972), and managers match situations to roles in a “logic of obligation” (March, 1982). It also finds that this sort of naturalistic decision-making is best modeled within domains, based on the kinds of decisions made in those domains (Barrett & Kurzban, 2006) rather than in generic, one-size-fits-all reasoning models. The logic and function of autobiographical narrative is, for instance, much more complicated than that of simple historical truth-discovery (Adler et al., 2017; Rasmussen & Habermas, 2011). Autobiographical reasoning has a directive function that involves understanding the past and predicting the future, but it also has important functions in self-definition and social engagement and functioning (Bluck et al., 2005). Work in all these domains is shot through with moral perception, evaluation, planning, and action. But this work has rarely been seen as work in moral psychology.

7.5 Reasoning Goes Awry: Two Paths to Moral Failure

One final thing needs to be said about moral reasoning: reason is often employed in the service of immoral and even evil goals. We mentioned this issue in Section 7.1, but here we catalog some of the ways that reasoning goes awry. This judgment that reasoning has gone bad is not based on rules of logic or other “normative” cognitive models. One can in fact closely follow rules of inference, deduction, and induction while planning genocide. Reason does not protect us from immorality. We investigate here some circumstances in which reason either transforms our perception of an evil goal into an acceptable or obligatory one or distances us from guilt.

Hart and colleagues (Hart, 2005b; Hart & Carlo, 2005) have used the term moral collapse to refer to moral failures like genocide. These instances are surely one end of a continuum of moral failure whose other end might begin with good deeds left undone. We will use the term *moral failure* here as a term of convenience to describe those instances ranging from participation in genocide to the omission of possible good. This allows us to look at the cognitive reasoning that contributes to these sorts of moral

failure. We will concentrate here on two paths that might lead one to bring the “hired gun” of reason into the service of moral failure: moral identity and empathy avoidance.

7.5.1 Moral Identity and Moral Failure

A variety of researchers have pointed to the centrality of moral identity in circumscribing the circle of moral concern (Aquino et al., 2006; Clayton & Opatow, 2003; Fowler, Law, & Gaesser, 2021; Hart, 2005a; Opatow, 2001, 2005, 2010; Skitka, 2010; Skitka et al., 2016; Skitka & Mullen, 2002). See Chapter 9, Section 9.3.1.3 for more detail on this. Once a group is excluded from moral concern, it becomes a legitimate, and possibly even obligatory, target for aggression. If the in-group is constituted in explicit rejection of an out-group that is perceived to be threatening (e.g. as in the case of war) then it may be morally obligatory to directly aggress against the out-group, to support in-group members who aggress, and to punish in-group members who do not. Moral identity need not be the source of such extreme behavior; one may simply ignore the needs of those outside the circle. This approach is a more profoundly social idea of moral identity than is often found in the literature.¹⁹ The point here is that once moral identity defines others as outside, then one can bring motivated reasoning (Epley & Gilovich, 2016; Klein & Epley, 2017) to bear to legitimize neglect or aggression.

7.5.2 Moral Distancing and Moral Failure

Another path that recruits cognitive strategies in service of moral failure is the avoidance of the aversive emotion that is aroused by watching others suffer. This suffering may not be directly or indirectly caused by the person watching. But the person watching may have a variety of reasons to avoid the aversive emotion: inability to escape the scene, a preference to pursue goals other than relieving the suffering, etc. In this case, cognitive distancing strategies can be called into service to reduce the aversive emotion.

Bandura (1999, 2002) provides a list of strategies that serve to produce psychological distance from the (in)action itself, the effects of the (in)action, from those harmed by action or inaction, or by reconstruing the nature of responsibility. The list attempts to be comprehensive but there is no theoretical reason to suppose it is completely exhaustive.²⁰ Some items from the list follow.

7.5.2.1 *Modifying the Descriptions of the Action*

- *Moral justification*: immoral conduct is made acceptable by portraying it as in service of a morally worthy goal.
- *Advantageous comparison*: comparing the current actions either to those they are reacting to (since Microsoft has done X, we can do Y) or to more severe possibilities (I know the data is insecure, but at least we are not selling it).

¹⁹Chapters 3, 5, and 9 explore this socially grounded moral identity in detail.

²⁰Our list includes most of the items listed by Bandura (1999, 2002) and several items from other authors where they seemed relevant.

- *Euphemistic labeling*: sanitizing (collateral damage of bombing), use of the passive (the workers became redundant), or use of specialized jargon (adverse consequences for medical accidents) to reduce awareness of the actual circumstances of the action.

7.5.2.2 *Modifying the Attitude Toward Consequences*

- *Minimizing*: altering the description of the consequences to make them seem less bad (e.g. undercounting casualties).
- *Ignoring*: various strategies will allow one to ignore consequences including physical distance (design of software is usually done at a distance) and redirection of attention (paying attention to other aspects of the situation). Those living near Auschwitz never thought to ask about the smoke, or if they did, they preferred not to know.
- *Misconstruing*: thinking that the outcome is different than it was. Often done by rapists but also by others (e.g. gassing as humane extermination).
- *Seeing injustice in nonmoral terms*: we can avoid seeing social injustice by seeing market forces or other “nonmoral” causes behind them. For instance, seeing job loss as simply a part of creative destruction in the economy rather than as a social problem that requires justice (Opatow, 2010).
- *Seeing injustice as misfortune*: seeing injustice as unfortunate luck or circumstances rather than as systematic societal arrangement of risk shunted onto certain categories of people. For instance displacement of poor neighborhoods for highways or crowding in inadequate housing (Opatow, 2010).

7.5.2.3 *Modifying Attitudes Toward the Victim*

- *Dehumanization*: making the victims of the action out to be less than human, and thus less deserving of moral consideration. Often done as a part of war. But also done in labels given computer users (e.g. lusers at MIT) and clients.
- *Blaming the victim*: claiming that the fate of the victim is in part caused by the victim’s own actions or characteristics (“If he weren’t so stupid I wouldn’t have to shock him” – from one of the participants in the Milgram obedience studies).

7.5.2.4 *Modifying Attitudes Toward Responsibility*

- *Displacement of responsibility*: claiming that someone or some circumstance other than the self is ultimately responsible. This is often done in hierarchy but also by consultants (one purpose of consulting is to acquire a detached party on whom the bad news can be blamed). In fact, almost anyone can find a way to blame either other people or circumstances that require them to act in morally reprehensible ways.
- *Diffusion of responsibility*: in most endeavors more than one person is required. When one acts as part of a group, the psychological burden of responsibility is lessened. When the group is responsible for a product, it is usually the case that individual members feel responsible only for those parts that directly concern them, and then only in relation to their role in the group (rather than their relation to the client/user).
- *Moral decoupling*: judging others’ or one’s own actions independently of their moral worth, e.g. praising a politician or sports star for skill while ignoring other moral failures (Bhattacharjee, Berman, & Reed, 2013).

- *Selective use of moral principles:* different moral principles (e.g. harm, justice, sacredness, etc.) can result in different evaluations of an action. Instead of a comprehensive analysis of moral principles, or even of those deemed most relevant, one can selectively choose those that support a desired evaluation. For instance, Uhlmann et al. (2009) found that Americans judging the killing of Iraqis by American forces were more likely to use reasoning that allowed for trading off costs and benefits than were those judging the killing of Americans by Iraqis.

This is an imposing list of the ways cognitive resources can be recruited in the service of moral failure. Of course, the list can be used for a positive purpose: to ask oneself or others if they are using these strategies without justification. The positive use of these strategies to check oneself or one's society is a crucial aspect of holding ourselves and others morally responsible, and of maintaining a cultural critique that influences how we think of our actions. So, a recognition of the tools that allow for bias based in cognitive reasoning can support moral critique.

7.5.3 Reflective Equilibrium and Moral Failure

This balancing of various considered moral judgments against each other was given a name by the philosopher Rawls (1971/1999): reflective equilibrium.²¹ We are using this method in our suggestion that recognizing the strategies of moral distancing and the limitations of moral identity can lead us to recognize and avoid moral failure. This equilibrium involves consulting those with expertise in an area, and also our intuitions (Daniels, 2018). In both this chapter and Chapter 9, we argue for the close integration of moral emotion in our moral intuition. And in Chapter 3 we document how closely linked rationality is to the specific moral ecologies in which it occurs. For this reason, we prefer explicitly to add social and emotional reflection to the idea of reflective equilibrium: social-emotional reflective equilibrium (Huff & Furchert, 2016). A considered balance of respected social influences, careful emotional responses, and conscious reflection can at least guard against moral failure, and is entirely in the spirit of the original proposal of reflective equilibrium. It is simply more explicit about what should be included in the equilibrium.

7.6 Discussion

7.6.1 Conclusion

We began this chapter with a reconsideration of the dominant paradigm in moral psychology for three decades, the cognitive developmental research program of Kohlberg. After reviewing the legacy of that program, we soon found ourselves in a web of different research programs that took us far from Kohlberg's commitment to consciously controlled pure reasoning processes and their relation to philosophical

²¹See Daniels (2018) for a history of the concept and its reception and critique and Narvaez (2010) for an integration of a similar concept into a psychological theory of moral expertise.

moral reasoning. This movement is reflected in the understandings we summarize here, and the questions that remain:

1. *We should be clearer what we mean when speaking of moral reason.*

The research conversation on cognition is changing in ways that depart dramatically from the old school of consciously controlled executive function. The move in cognitive psychology toward implicit cognition, two- (or more) process models of cognition, the incorporation of emotion, and naturalistic cognition should lead us away from a normative commitment to conscious rational processing and toward a descriptive understanding of the ways that different kinds of processes we call reasoning are involved in moral action. When speaking of reason and moral action, we will need to be more specific about what type of reason, what capabilities that type has, the biases associated with it, and how it interacts with other systems.

2. *There is a great deal of moral cognition that is not conscious.*

The Kohlberg research tradition in moral psychology championed an emphasis on conscious logical thinking as the foundational aspect of morality. But as cognitive psychology itself has opened its doors to the varieties of cognition in human living, moral psychology is also beginning to open its doors and recognize the many ways cognitive processing interacts with moral life. Many cognitive processes are automatic – either in origin (e.g. perceptual judgment) or because they are so well practiced, they require little working memory (e.g. expertise-based judgment). Moral exemplars, people who excel at particular kinds of moral action, report expending little of their energy on deciding what is right or wrong and the most energy on planning and implementing their goal-directed activity. So, even in this chapter on cognitive processes – or perhaps especially in this chapter on cognitive processes – we find a shift away from the role of conscious deliberation and toward the multiplicity of ways that cognition is “the servant of” and even the “hired gun for” (im)moral action.

3. *Moral cognition is for moral acting.*

This is an application of Fiske’s (1992) admonition that “thinking is for doing.”²² It recognizes that much moral action is based in instrumental reasoning of how to achieve moral goals that have already been adopted. And that the skills of moral perception, moral reflection, self-regulation etc. that are called upon in moral action (see Chapter 6) have a rationality to them that may be implicit and domain specific. This gives reason a greater role than the biased motivated reasoning in the “lawyer” metaphor in the new synthesis proposal (Haidt, 2007). Surely, we use motivated reasoning to simply buttress our own positions or defend our prejudices. But when people do search for the right thing to do, they often use cognitive processes (in interaction with emotion and social identity) to critique their own culture and plan action to change it (Killen & Dahl, 2021). Thus, when considered as an analysis of cultural critique, much of the original work in the Kohlberg tradition still stands and can provide important structure to further investigation. But much more work remains to be done in how these conscious

²²Which she takes from William James (1890/1950, p. 959) *Principles of Psychology*.

processes interact with automatic processes and other influences to structure and guide moral action. We find people constantly choosing their way among the multiple influences on their (im)moral goals and plans of action. The change in focus, then, is from isolated conscious moral problem-solving of self-contained moral puzzles and toward goal-directed action and all the cognitive processes that support the goals and the action.

4. *Moral reason can be both a useful guide and a source of bias.*

One of the central motivators of the insistence on conscious moral cognition as the “real” moral cognition (whether in Kohlberg or in two-process models) is the idea that only principled reasoning can save us from personal bias and relativism. There are several difficulties with this line of approach to moral cognition. First, the role of reason is already morally ambiguous. Clearly, reason has a strong role to play in *justifying* or *protecting* moral failure through a variety of moral distancing mechanisms. Second, the reduction of morality to a simple power struggle is not the only possible option after one opens the door to other influences on moral cognition. Many theorists allow that a critique of moral culture can be more than a simple power struggle. It can include appeals to intuition, to an objective moral charter, to shared humanity, to religion, or to any of multiple values that can balance each other in the moral ecology of a culture. Of course, this balance is often lost, resulting in moral collapse. One of the themes of Chapter 3 is the search for how a moral ecology might maintain an appropriate social-emotional reflective equilibrium. The final problem in viewing reason as the bulwark against relativism and moral collapse is that moral collapse most often happens because of an imbalance in the moral ecology, when, for example, moral identity is too closely tied to the interests of an in-group. Rationality does not guard against such dangers, instead it can become the “hired gun” that serves the imbalance.

7.6.2 Application

At the end of Chapter 3, you can read about the shifting moral ecologies associated with Nathaniel Borenstein’s decision to become a NATO consultant. Because of his expertise in online learning systems, and despite his history as a peace activist, he had been recruited by NATO to help them design a training system for tactical nuclear missile launchers. While doing so, he found commercial firms making evidence-based arguments that including modern training software and help systems on missile launchers was the best way to upgrade the skills of military personnel who operated the launchers. These arguments were rational, plausible, and based in psychological evidence. But due to his experience developing software for real systems, Borenstein was skeptical. Consider this quote from Borenstein (1989, 17):

Solving one problem often creates a host of new ones. I devote my efforts rather narrowly to learning how to make computers easier for people to use. But more usable programs are not always better in every sense. Making a program seem simple to the user usually means making its internals far more complex, and this often makes it less reliable. This is not the kind of tradeoff one makes lightly when dealing with computers that control nuclear weapons.

This is ethical rationality based in expertise in a domain. The reasoning here is *not* high-level ethical argument about the good. Borenstein already knows his moral commitments – he is an unalloyed pacifist. But as a software developer, he knows that all design choices involve value tradeoffs. And he has extensive experience with one particular one: trading off reliability for ease of use in the human interface. He recognizes the tradeoff and its dangers easily, almost automatically and without conscious search. He reports an emotional reaction (he is frightened by the tradeoff) based on his appraisal of the context (safety-critical weapons systems). The emotion is part of what convinces him to agree to a consulting position. He was influenced by a phone call from “a pleasant-sounding” lieutenant commander with a military moral ecology he perceived as being in opposition to his moral commitments (see Chapter 3). His participation in the panel resulted in a less dangerous outcome in software design and he reports “I was entirely unprepared to like the military men as much as I did. They were smart, conscientious, and peaceable” (Borenstein 1989, 18).

Borenstein’s turn in military consulting provides us with an example of the complexities of moral reason in real-life situations. There is a great deal of cognitive processing, mixed with emotional responses and automatic, almost intuitive, reactions. There is very little abstract thinking about moral principles. The moral reasoning and moral action are deeply interwoven into expertise in a particular domain (software development) and embedded in multiple complex moral ecologies.

7.6.3 Open Questions

1. *Where in human living does moral evaluation occur and what is it for?*

The move away from explicit cognition in moral psychology has occurred in parallel to a similar move in cognitive psychology. Moral psychologists have begun to adopt a multiplicity of descriptive and instrumentalist approaches, including evolutionary, two-process, expertise-based, and naturalistic moral judgment and action models. These models have expanded our view of the role of reason in moral action in the same way that they have expanded our understanding of rationality in general. First, instead of focusing on isolated judgment, they are much more likely to focus on goal-directed action and the role cognitive processing plays in it. This change in focus is a central aspect of this book: it argues for the study of moral *action* and the influences and processes that support it. This is parallel to the instrumentalist approaches in human cognition. These approaches ask, “What is cognitive processing for?” The instrumentalist approach in moral psychology asks also, “How is cognitive processing involved in moral action?”

Similarly, the descriptive approaches to cognition ask, “Where in human living is cognitive processing done?” Moral expertise approaches ask a similar question, as do anthropological models of moral cognition, as does the extensive literature on moral distancing. They thus expand our vision of where and when moral cognition operates, bringing into view a wider span of human action. The works of Malle (2021), Killen and Dahl (2021), and Epley and Tanenbaum (2017) are examples of promising steps in this direction.

2. *Is it necessary to resolve the reason-emotion tension in moral psychology?*

In short, no. The proponents on the two sides in the moral reason vs. emotion debate are in essential agreement that both matter. The exchanges frame the

disagreement more as a differential weighting of reason vs. intuition/emotion. But this is too abstract. It might well be that reason plays a predominant role in some aspects of moral action (e.g. cultural critique, planning) and intuition/emotion plays a dominant role in others (e.g. compassion, commitment to social change), and that in most all moral action there is some level of balance. How these balances are achieved is the subject of the next question.

3. *Might it be better to study moral evaluation rather moral reasoning?*

When an author wants to appeal to reason rather than bias, *reason* is often the preferred term (rather than *cognitive processing*). But even cognitive processing assumes that only cognition is what is under investigation. There is little neurological or psychological warrant for this assumption. We might better use *moral evaluation* to include a range of influences when cognitive, emotional, and social evaluation is occurring in the service of moral action. This avoids falling back on the rationalist, foundationalist tropes of earlier approaches. And it allows for something like *social-emotional reflective equilibrium* as a process whose interactions, influences, processes, and limitations might be investigated.

7.7 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- Bandura (2002). “Selective moral disengagement in the exercise of moral agency.” A systematic review of the ways that motivated reasoning can distance us psychologically from the evil we do.
- Evans and Elqayam (2011). “Towards a descriptivist psychology of reasoning and decision making.” A useful overview of the question “what is reason?” in the cognitive psychology literature.
- Gibbs et al. (2007). “Moral judgment development across cultures: Revisiting Kohlberg’s universality claims.” A relatively comprehensive review of the critiques of the Kohlberg system and the empirical literature supporting those critiques.
- Killen and Dahl (2021). “Moral reasoning enables developmental and social change.” A review of the parallels in moral reasoning that lead to individual development and to social change. Presents moral reasoning development in the light of lived experience.
- Kohlberg (1971). “From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development.” The classic statement of how the Kohlberg system moves back and forth from descriptive “is” statements and normative “ought” statements.
- Malle (2021). “Moral Judgments.” A rearranging of the pieces in moral judgment that orients us more toward a descriptive psychology of types of moral judgment and their contexts.
- Rest et al. (1999). “A Neo-Kohlbergian approach: The DIT and schema theory.” The initial presentation of “neo” Kohlberg from the research group that designed and developed the most widely used test of Kohlbergian moral reasoning.

References

- Adler, J. M., Dunlop, W. L., Fivush, R. et al. (2017). Research methods for studying narrative identity. *Social Psychological and Personality Science*, 8(5), 519–527. doi:10.1177/1948550617698202
- Aktipis, C. A., & Kurzban, R. O. (2004). Is homo economicus extinct? Vernon Smith, Daniel Kahneman and the evolutionary perspective. *Advances in Austrian Economics*, vol. 7. London: Emerald
- Aquino, K., Reed, A., Thau, S., & Freeman, D. (2006). A grotesque and dark beauty: How moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. *Journal of Experimental Social Psychology*, 43, 385–392. doi:10.1016/j.jesp.2006.05.013
- Augustine. (397/2007). *Confessions* (E. J. Sheed, Trans.). Indianapolis, IN: Hackett.
- Bago, B., & De Neys, W. (2019). The Smart System 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, 25(3), 257–299. doi:10.1080/13546783.2018.1507949
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, 31(2), 101–119.
- Bardi, A., & Schwartz, S. H. (2003). Values and behavior: strength and structure of relations. *Personality and Social Psychology Bulletin*, 29(10), 1207–1220. doi:10.1177/0146167203254602
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.
- Barrett, H. C., & Kurzban, R. O. (2006). Modularity in cognition: framing the debate. *Psychological Review*, 113(3), 628–647. doi:10.1037/0033-295X.113.3.628
- Barsolou, L. W., Simmons, W. K., Barbey, A. K. et al. (2003). Grounding conceptual knowledge in modality specific systems. *Trends in Cognitive Sciences*, 7(2), 84–91.
- Bellah, R. N. (2011). *Religion in Human Evolution: From the Paleolithic to the Axial Age*. Cambridge, MA: Belknap.
- Beller, S. (2010). Deontic reasoning reviewed: psychological questions, empirical findings, and current theories. *Cognitive Processing*, 11(2), 123–132.
- Bhattacharjee, A., Berman, J. Z., & Reed, A. (2013). Tip of the hat, wag of the finger: How moral decoupling enables consumers to admire and admonish. *Journal of Consumer Research*, 39(6), 1167–1184.
- Bluck, S., Alea, N., Habermas, T. et al. (2005). A tale of three functions: The self-reported uses of autobiographical memory. *Social Cognition*, 23(1), 91–117.
- Bok, H. (2014). Baron de Montesquieu, Charles-Louis de Secondat. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/montesquieu/>
- Borenstein, N. S. (1989). My life as a NATO collaborator. *Bulletin of the Atomic Scientists*, April, 14–20.
- Borg, I., Bardi, A., & Schwartz, S. H. (2015). Does the value circle exist within persons or only across persons? *Journal of Personality*, 85(2). doi:10.1111/jopy.12228
- Breuer, J., & Freud, S. S. (1895/2000). *Studies in Hysteria* (J. Strachey & A. Freud, Trans.).
- Cesario, J., Johnson, D. J., & Eisthen, H. L. (2020). Your brain is not an onion with a tiny reptile inside. *Current Directions in Psychological Science*, 29(3), 255–260. doi:10.1177/0963721420917687
- Chater, N., & Oaksford, M. (1996). Deontic reasoning, modules and innateness: A second look. *Mind & Language*, 11(2), 191–202.

- Clayton, S., & Opatow, S. (2003). Justice and identity: Changing perspectives on what is fair. *Personality and Social Psychology Review*, 7(4), 298–310. doi:10.1207/S15327957PSPR0704_03
- Colby, A., Kohlberg, L., Gibbs, J. C. et al. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development*, 48(1/2), 1–124.
- Cosmides, L., & Tooby, J. (2008). Can a general deontic logic capture the facts of human moral reasoning? How the mind interprets social exchange rules and detects cheaters. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 1, pp. 53–119). Cambridge, MA: MIT Press.
- Cushman, F., Young, L., & Hauser, M. D. (2006). The role of conscious reasoning and intuition in moral judgment. *Psychological Science*, 17(12), 1082–1089.
- Daniels, N. (2018). Reflective equilibrium. In E. N. Zalta & U. Nodelman (Eds.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/ENTRIES/reflective-equilibrium/>
- De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, 7(1), 28–38. doi:10.1177/1745691611429354
- De Neys, W. (2021). On dual- and single-process models of thinking. *Perspectives on Psychological Science*, 16(6), 1412–1427. doi:10.1177/1745691620964172
- De Silva, C., & Habermas, T. (2011). Narrative means to manage responsibility in life narratives across adolescence. *The Journal of Genetic Psychology*, 172(1), 1–20.
- Dewey, J., & Tufts, J. H. (1908). *Ethics*. New York: Holt.
- Dinh, J. E., & Lord, R. G. (2013). Current trends in moral research: What we know and where to go from here. *Current Directions in Psychological Science*, 22(5), 380–385. doi:10.1177/0963721413486147
- Dreyfus, H. L., & Dreyfus, S. E. (2004). The ethical implications of the five-stage skill-acquisition model. *Bulletin of Science, Technology, & Society*, 24, 251–264. doi:10.1177/0270467604265023
- Elqayam, S., & Evans, J. S. B. T. (2011). Subtracting “ought” from “is”: Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, 34(05), 233–248. doi:10.1017/s0140525x1100001x
- Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(3), 133–140. doi:10.1257/jep.30.3.133
- Epley, N., & Tannenbaum, D. (2017). Treating ethics as a design problem. *Behavioral Science & Policy*, 3(2), 73–84.
- Evans, J. S. B. T., & Elqayam, S. (2011). Towards a descriptivist psychology of reasoning and decision making. *Behavioral and Brain Sciences*, 34(5), 275–290. doi:10.1017/S0140525X1100001X
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3), 223–241. doi:10.1177/1745691612460685
- Fischhoff, B., & Downs, J. S. (1997). Accentuate the relevant. *Psychological Science*, 8(3), 1–5.
- Fiske, A. P. (2002). Using individualism and collectivism to compare cultures – A critique of the validity and measurement of the constructs: Comment on Oyserman et al. (2002). *Psychological Bulletin*, 128(1), 78–88. doi:10.1037//0033-2909.128.1.78
- Fiske, A. P., & Haslam, N. (2005). The four basic social bonds: Structures for coordinating interaction. In M. Baldwin (Ed.), *Interpersonal Cognition* (pp. 267–298). New York: Guilford Press.
- Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from Daguerreotype to laser photo. *Journal of Personality and Social Psychology*, 63(6), 877–889. doi:10.1037/0022-3514.63.6.877
- Fodor, J. (1983). *The Modularity of Mind*. Scranton, PA: Crowell.
- Fowler, Z., Law, K. F., & Gaesser, B. (2021). Against empathy bias: The moral value of equitable empathy. *Psychological Science*, 32(5), 766–779. doi:10.1177/0956797620979965

- Frankena, W. (1939). The naturalistic fallacy. *Mind*, 48, 464–477.
- Frimer, J. A., Tell, C. E., & Haidt, J. (2015). Liberals Condemn Sacrilege Too: The Harmless Desecration of Cerro Torre. *Social Psychological and Personality Science*, 6(8), 878–886. doi:10.1177/1948550615597974
- Frimer, J. A., & Walker, L. J. (2008). Towards a new paradigm of moral personhood. *Journal of Moral Education. Special Issue: Towards an Integrated Model of Moral Functioning*, 37(3), 333–356. doi:10.1080/03057240802227494
- Galotti, K. M. (2007). Decision structuring in important real-life choices. *Psychological Science*, 18(4), 320–325.
- Gibbs, J. C. (1977). Kohlberg's stages of moral judgment: A constructive critique. *Harvard Educational Review*, 47, 43–61.
- Gibbs, J. C., Basinger, K. S., Grime, R. L. et al. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review*, 27(4), 443–500. doi:10.1016/j.dr.2007.04.001
- Gilligan, C. (1977). In a different voice: Women's conceptions of self and morality. *Harvard Educational Review*, 47(4), 481–517.
- Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In P. R. Shaver & M. Mikulincer (Eds.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil* (pp. 11–31). New York: American Psychological Association Books.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029–1046.
- Graham, J., Nosek, B. A., Haidt, J. et al. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101, 366–385.
- Gregg, G. S. (2005). *The Middle East: A Cultural Psychology*. New York: Oxford University Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998.
- Haidt, J., Björklund, F., & Murphy, S. (2000). *Moral Dumbfounding: When Intuition Finds no Reason*. University of Virginia. Charlottesville, VA.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, 65(4), 613–613.
- Hart, D. (2005a). Adding identity to the moral domain. *Human Development*, 48(4), 257–261.
- Hart, D. (2005b). The development of moral identity. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (pp. 166–196). Lincoln, NE: University of Nebraska Press.
- Hart, D., & Carlo, G. (2005). Moral development in adolescence. *Journal of Research on Adolescence. Special Issue: Moral Development*, 15(3), 223–233. doi:10.1111/j.1532-7795.2005.00094.x
- Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science*, 8(2), 195–207. doi:10.1177/1745691612460684
- Hassin, R. R., Bargh, J. A., & Zimmerman, S. (2009). Automatic and flexible: The case of nonconscious goal pursuit. *Social Cognition*, 27(1), 20–36.
- Helzer, E. G., & Pizarro, D. A. (2011). Dirty liberals! Reminders of physical cleanliness influence moral and political attitudes. *Psychological Science*, 22(4), 517–522.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. Thousand Oaks, CA: Sage.

- Huff, C. (2014). From meaning well to doing well: Ethical expertise in the GIS domain. *Journal of Geography in Higher Education*. doi:10.1080/03098265.2014.936314
- Huff, C., & Furchert, A. (2016). Embracing the complexity of moral action. In D. Bertini (Ed.), *Moral Heteronomy. History, Proposals, Arguments, a special issue of Dialegetha*.
- Hume, D. (1739/1985). *A Treatise of Human Nature*. New York: Penguin Books.
- Inglehart, R., & Welzel, C. (2005). *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.
- Iyer, R., Koleva, S., Graham, J. et al. (2012). Understanding libertarian morality: the psychological dispositions of self-identified libertarians. *Plos One*, 7(8), e42366. doi:10.1371/journal.pone.0042366
- James, W. (1890/1950). *The Principles of Psychology*. Mineola, NY: Dover.
- Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape: moral motives and group-based moralities. *Personality and social psychology review*, 17(3), 219–236. doi:10.1177/1088868313480274
- Janoff-Bulman, R., & Carnes, N. C. (2018). The model of moral motives: A map of the moral domain. In K. Gray & J. Graham (Eds.), *Atlas of Moral Psychology* (pp. 223–232). New York: Guilford Press.
- Janoff-Bulman, R., Sheikh, S., & Baldacci, K. G. (2008). Mapping moral motives: Approach, avoidance, and political orientation. *Journal of Experimental Social Psychology*, 44(4), 1091–1099.
- Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: two faces of moral regulation. *Journal of Personality and Social Psychology*, 96(3), 521–537. doi:10.1037/a0013779
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Keane, W. (2015). *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton University Press.
- Kierkegaard, S. (1849/2004). *The Sickness Unto Death* (A. Hannay, Trans.). New York: Penguin.
- Killen, M., & Dahl, A. (2021). Moral reasoning enables developmental and social change. *Perspectives on Psychological Science*, 16(6), 1209–1225. doi:10.1177/1745691620964076
- Klein, N., & Epley, N. (2017). Less evil than you: Bounded self-righteousness in character inferences, emotional reactions, and behavioral extremes. *Personality and Social Psychology Bulletin*, 43(8), 1202–1212. doi:10.1177/0146167217711918
- Kohlberg, L. (1958). The development of modes of moral thinking and choice in the years 10 to 16. PhD Dissertation. University of Chicago.
- Kohlberg, L. (1964). The development of moral character and moral ideology. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of Child Development Research* (Vol. 1, pp. 383–431). New York: Russell Sage Foundation.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive Development and Epistemology* (pp. 151–235). New York: Academic Press.
- Kohlberg, L., Kramer, R. M. C., & Karen S. (1969). Continuities and discontinuities in childhood and adult moral development. *Human Development*, 12, 93–120.
- Kohlberg, L., Levine, C., & Hewer, A. (1983). *Moral Stages: A Current Formulation and a Response to Critics*. Basel: S. Karger.
- Kohlberg, L., & Power, F. C. (1981). Moral development, religious thinking, and the question of a seventh stage. *Zygon*, 16(3), 203–259.
- Korsgaard, C. M. (2006). Morality and the distinctiveness of human action. In S. Macedo & J. Ober (Eds.), *Primates and Philosophers: How Morality Evolved* (pp. 98–119). Princeton, NJ: Princeton University Press.
- Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, 10(2), 83–109.

- Lagattuta, K. H., & Weller, D. (2014). Interrelations between theory of mind and morality: A developmental perspective. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (2nd ed., pp. 385–407). New York: Psychology Press.
- Levine, C., Kohlberg, L., & Hewer, A. (1985). The current formulation of Kohlberg's theory and a response to critics. *Human Development*, 28, 94–100.
- Lipschitz, R., Klein, G., Orasanu, J. et al. (2001). Focus article: Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, 14(5), 329–389.
- Malle, B. F. (2021). Moral Judgments. *Annual Review of Psychology*, 72, 293–318. doi:10.1146/annurev-psych-072220-104358
- Malti, T., & Ongley, S. F. (2014). The development of moral emotions and moral reasoning. In M. Killen & J. Smetana (Ed.), *Handbook of Moral Development* (2nd ed., pp. 163–183). New York: Psychology Press.
- March, J. G. (1982). Theories of making choice and making of decisions. In R. Wolf (Ed.), *Organizing Industrial Development* (pp. 305–326). Munich: De Gruyter.
- McGrath, S. J. (2012). *The Dark Ground of the Spirit: Schelling and the Unconscious*. New York: Routledge/Taylor Francis.
- Moore, G. E. (1903). *Principia Ethica*. New York: Cambridge University Press.
- Narvaez, D. (2005). The neo-Kohlbergian tradition and beyond: Schemas, expertise, and character. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (Vol. 51, pp. 119–163). Lincoln, NE: University of Nebraska Press.
- Narvaez, D. (2010). Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Narvaez, D., & Bock, T. (2002). Moral schemas and tacit judgement or how the Defining Issues Test is supported by cognitive science. *Journal of Moral Education*, 31(3), 297–314.
- Narvaez, D., & Bock, T. (2014). Developing ethical expertise and moral personalities. In L. P. Nucci, D. Narvaez, & T. Krettenauer (Eds.), *Handbook of Moral and Character Education* (pp. 140–158). New York: Routledge.
- Narvaez, D., & Lapsley, D. K. (2005). The psychological foundations of everyday morality and moral expertise. In D. Lapsley & F. C. Power (Eds.), *Character Psychology and Character Education* (pp. 140–165). South Bend, IN: University of Notre Dame Press.
- Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.
- Nisbett, R. E., & Ross, L. (1980). *Human Inference: Strategies and Shortcoming of Social Judgment*. Mahwah, NJ: Prentice Hall.
- Opatow, S. (2001). Social injustice. In D. J. Christie, R. V. Wagner, & D. A. Winter (Eds.), *Peace, Conflict, and Violence: Peace Psychology for the 21st Century*. Englewood Cliffs, NJ: Prentice-Hall.
- Opatow, S. (2005). Moral exclusion to moral inclusion: Theory for teaching peace. *Theory into Practice*, 44(4), 303–318.
- Opatow, S. (2010). Social injustice and moral exclusion. In N. J. Young (Ed.), *The Oxford International Encyclopedia of Peace*. New York: Oxford.
- Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin and Review*, 11(6), 988–1010.
- Over, D. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 3–18). Hoboken, NJ: Blackwell.
- Payne, J. W., & Bettman, J. R. (2004). Walking with scarecrow: The information-processing approach to decision research. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 111–132). Hoboken, NJ: Blackwell.

- Pietraszewski, D., & Wertz, A. E. (2022). Why evolutionary psychology should abandon modularity. *Perspectives on Psychological Science*, 17(2), 465–490 doi:10.1177/1745691621997113
- Pincoffs, E. (1971). Quandry Ethics. *Mind*, 80(320), 552–571.
- Prinz, J. J. (2008). Is morality innate? In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 1, pp. 367–406). Cambridge, MA: MIT Press.
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology Is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, 118(1), 57–75.
- Rasmussen, A. S., & Habermas, T. (2011). Factor structure of overall autobiographical memory usage: The directive, self, and social functions revisited. *Memory*, 19(6), 597–605.
- Rawls, J. (1971/1999). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rest, J. R., Narvaez, D., Bebeau, M. J. et al. (1999). A Neo-Kohlbergian approach: The DIT and schema theory. *Educational Psychology Review*, 11(4), 291–324.
- Rest, J. R., Narvaez, D., Thoma, S. J. et al. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology*, 91(4), 644–659.
- Richards, P. S., & Davison, M. L. (1992). Religious bias in moral development research: A psychometric investigation. *Journal for the Scientific Study of Religion*, 31(4), 467–485.
- Roccas, S., Sagiv, L., Schwartz, S. H. et al. (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin*, 28(6), 789–801.
- Rozin, P., Lowery, L., Imada, S. et al. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76(4), 574–586.
- Sayre-McCord, G. (2015). Moral realism. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/value-theory/>
- Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, 41(8), 1147–1163 doi:10.1177/0146167215591501
- Schelling, F. W. J. (1800/1993). *System of Transcendental Idealism* (P. Heath, Trans.). Charlottesville, VA: University of Virginia Press.
- Schön, D., A. (1984). *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- Schroeder, M. (2021). Value theory. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. <https://plato.stanford.edu/entries/value-theory/>
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 1–65). New York: Academic Press.
- Schwartz, S. H. (2006). Basic human values: Theory, measurement and applications. *Revue française de sociologie*, 4(47), 929–968.
- Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial Motives Emotions, and Behavior: The Better Angels of Our Nature* (pp. 221–241). Washington, DC: American Psychological Association.
- Schwartz, S. H. (2016). Basic individual values: Sources and consequences. In D. Sander & T. Brisch (Eds.), *Handbook of Value* (pp. 63–84). Oxford: Oxford University Press.
- Schwartz, S. H., & Bilsky, W. (1987). Toward a universal structure of human values. *Journal of Personality and Social Psychology*, 53(3), 550–562.
- Schwartz, S. H., Cieciuch, J., Vecchione, M. et al. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, 103(4), 663–688.
- Searle, J. R. (1964). How to derive “ought” from “is”. *Philosophical Review*, 73(1), 43–58.
- Shweder, R. A., Mahapatra, M., & Miller, J. G. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The Emergence of Morality in Young Children* (pp. 1–83). Chicago: University of Chicago Press.

- Shweder, R. A., & Menon, U. (2014). Old questions for the new anthropology of morality: a commentary. *Anthropological Theory*, 14(3), 356–370.
- Shweder, R. A., Much, N. C., Mahapatra, M. et al. (1997). The “Big Three” of morality (autonomy, community, and divinity) and the “big three” Explanations of Suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and Health* (pp. 119–169). New York: Routledge.
- Simon, H. A. (1983). *Reason in Human Affairs*. Stanford, CA: Stanford University Press.
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, 4(4), 267–281.
- Skitka, L. J., Hanson, B. E., & Wisneski, D. C. (2016). Utopian hopes or dystopian fears? Exploring the motivational underpinnings of moralized political engagement. *Personality and Social Psychology Bulletin*, 32(2), 177–190. doi:10.1177/0146167216678858
- Skitka, L. J., & Morgan, G. S. (2014). The social and political implications of moral conviction. *Political Psychology*, 35, 95–110. doi:10.1111/pops.12166
- Skitka, L. J., & Mullen, E. (2002). The dark side of moral conviction. *Analyses of Social Issues and Public Policy*, 2(1), 35–41.
- Smith, A. (1759/2009). *The Theory of Moral Sentiments*. New York: Penguin.
- Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, 28(04), 531–542.
- Thoma, S. J. (2014). Measuring moral judgments from a Neo-Kohlbergian perspective. Paper presented at Can Virtue Be Measured? Oxford, UK.
- Thoma, S. J., Derryberry, W. P., & Crowson, H. M. (2013). Describing and testing an intermediate concept measure of adolescent moral thinking. *European Journal of Developmental Psychology*, 10(2), 239–252. doi:10.1080/17405629.2013.766129
- Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology*, 64, 231–255. doi:10.1146/annurev-psych-113011-143812
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Uhlmann, E. L., Pizarro, D. A., Tannenbaum, D. et al. (2009). The motivated use of moral principles. *Judgment and Decision Making*, 4(6), 476–491.
- Van de Vondervoort, J. W., & Hamlin, J. K. (2018). The early emergence of sociomoral evaluation: infants prefer prosocial others. *Current Opinion in Psychology*, 20, 77–81. doi:10.1016/j.copsyc.2017.08.014
- Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. *Child Development*, 55(3), 667–691.
- Walker, L. J., & Frimer, J. A. (2008). Being good for goodness’ sake: Transcendence in the lives of moral heroes. In F. Oser & W. Veugelers (Eds.), *Getting Involved: Global Citizenship Development and Sources of Moral Values* (pp. 309–326). Rotterdam: Sense.
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, 86(4), 629–647. doi:10.1037/0022-3514.86.4.629
- Walsh, C. (2001, October 1). Reconstructing Larry: Assessing the legacy of Lawrence Kohlberg. *HGSE News*. Retrieved from http://www.gse.harvard.edu/news/features/larry10012000_page1.html
- Wason, P. C., & Evans, J. S. B. T. (1975). Dual processes in reasoning? *Cognition*, 3(2), 141–154.
- Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, 60, 53–85.
- Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, 15, 728–755. doi:10.1080/17405629.2018.1435413
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16, 780–784.

- Wynn, K., & Bloom, P. (2014). The moral baby. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Development* (pp. 435–453). New York: Psychology Press.
- Yamagishi, T., Li, Y., Takagishi, H. et al. (2014). In search of homo economicus. *Psychological Science*, 25(9), 1699–1711. doi:10.1177/0956797614538065
- Zapko-Willmes, A., Schwartz, S. H., Richter, J. et al. (2021). Basic value orientations and moral foundations: convergent or discriminate constructs? *Journal of Research in Personality*, 92, 104099. doi:https://doi.org/10.1016/j.jrp.2021.104099
- Zsombok, C. E., & Klein, G. (Eds.). (2014). *Naturalistic Decision Making*. New York: Psychology Press.

Final Proofs

Moral Emotion

What are emotions, how do they work, and why do they not want us to achieve our best?

(Cohen, 2005, p. 5)

Emotions are meaning.

(Barrett, 2017, p. 126)

There is nothing more ordinary than to talk of the opposition of reason and emotion. It has been a theme in Eastern and Western philosophy since earliest times (Solomon, 2002) and continues until this day. In the first quotation from Cohen, we have a classic presentation of emotions as the enemy, or at least as unruly housemates we do not understand. In this chapter, we will attempt to convince you that things are more complicated than this simple, though popular, dualism – that the role of emotions in our moral lives is crucial and complex – and that emotions can sustain that role because they “are meaning,” or one might even say, they are the heart of meaning.

In Chapter 7, we began with a short overview of reason itself. We will follow the same practice here and ask with Cohen, “What are emotions, how do they work?” We will move from there to considering the role of emotions in morality and what have been called the “moral emotions.” Then we will consider whether emotions are something beyond our control or whether we can change or train our emotional responses. And we will conclude by arguing that (1) we could not be moral without being emotional and (2) it is time to think about how one would train the moral emotions.

8.1 An Overview of Theories of Emotion

One might expect a chapter like this to offer a high-level theory of the emotions. This cannot easily be done well. Some theorists have even suggested that a “general theory [of emotions] is neither necessary nor possible” (Nesse, 2014, p. 321). The

deceptive single word emotion refers to an experience that varies dramatically in intensity, time course, physiological expression, action tendency, cognition, and of course, culture. Theories of emotion often succeed either by ruling out significant and interesting phenomena or by retreating to a low-level reduction that loses touch with the actual experience and complexity of the phenomenon. Thankfully, there is at least some structure in the kinds of theoretical approaches. We present it here in terms of four different categories of theories – *basic emotion theories*, *appraisal theories*, *cultural theories*, and *constructivist theories* (Cameron, Lindquist, & Gray, 2015) – with each category allowing for somewhat greater complexity in description of the process and experience.¹ Thus, as in many chapters in this book, we begin in principle without a definition, in order to expand our horizon and include potentially interesting phenomena.

8.1.1 Basic Emotion Theories

Most introductory texts in psychology begin their section on emotion by talking about theories of “basic emotion.” This has the helpful pedagogical effect that it connects to the evolution and neuroscience chapters, and that it treats emotions as the separate things they often appear to be in language and in common sense. It also has distinct disadvantages: it treats emotions as wholly separate things, perpetuates the reason–emotion distinction, and only includes culture as an afterthought.

What the basic emotion theories share is (Cameron et al., 2015; Ekman, 1999; Ekman & Cordaro, 2011; Tracy & Randles, 2011):

- a conception of “basic” as grounded in a specific *physiological* underpinning²
- a grounding in an *evolutionary* background, with the most basic emotions appearing in some form in other species
- the prediction of strong *cross-cultural similarities* in elicitors, expression, and action motivation

¹Note that there are even disagreements about the appropriate number of categories for the kinds of emotion theories. Ellsworth (2013) lists five categories: adaptive, basic emotions, dimensional, appraisal, and constructivist theories. Scherer and Peper (2001) list seven categories: adaptational models, dimensional models, motivational models, circuit and discrete emotion models, and meaning and construct models. Barrett et al. (2019) provide a two-dimensional model of the conceptual space that emotion theories inhabit. The category system you choose depends on what differences and similarities you want to emphasize.

²See Solomon (2002) for how this choice of the meaning of “basic” is already full of assumptions. One alternative for basic could be “basic in human experience,” such as anxiety for Kierkegaard (1849/2004) or compassion (sympathy) for Hume (1739/1985) or lust(desire) for Freud (1905/1962). Another alternative is “basic to human sociality,” such as anger for Aristotle (Sokolon, 2006). Ancient Hindu culture produced its own listing of basic emotions rooted in the role they play in poetry, drama, and moral life (Shweder et al., 2008).

Most of the researchers in this tradition have been engaged in providing evidence of these characteristics for that subset of emotions that they list as basic.³ The huge complexity and variety of emotions is then produced by combining these elements or by rapid fluctuations between two or more of the basic emotions (Ekman & Cordaro, 2011; Ortony & Turner, 1990; Vaccaro, Kaplan, & Damasio, 2020). Basic emotions can also be modified by learning so that elicitors, expression, and action motivation are changed and shape the emotional process (see, e.g. work on display rules). If I am carrying a large grocery bag home, I might react with surprise if it splits open to spill my groceries on the sidewalk. This would elicit a characteristic body posture and facial expression, which might immediately be overridden by frantic attempts to rescue things, themselves overridden by attempts to display calm competence or humor. The mixed emotions I experience, and the manner in which I gather my groceries, would be shaped by which elicitors were present in that situation (e.g. other people watching).

Because these approaches ground emotion in evolutionarily shaped physiological responses to the environment, basic emotion theories are also those that are most interested in establishing the causal chain of how “an emotion” proceeds.⁴ Thus, some versions see the physiological responses of an emotion as mostly automatic and prior to any thought, with thought chasing behind the already unfolding emotion and giving reasons for it. Simplistic versions of this end up justifying the emotion–reason dichotomy. More complex versions of basic emotion theories are the prototype for those approaches to moral emotions (Rozin et al., 1999) that see a limited number of basic moral emotions with identifiable elicitors (injustice), expressions (lowered eyebrow stare), and action motivations (assertion or aggression).

8.1.2 Appraisal Emotion Theories

Appraisal theories see emotion experiences as structured appraisals of and motivations toward the world. Appraisal theories come in a variety of flavors, but they all share the mixing of cognition and emotion into the overall experience and shape of the emotion. The entire process is conceived as a dense weave of feedback loops among conscious and nonconscious cognition, physiology, subjective feel, action motivation, and action.

The central process, not surprisingly, is *appraisal*, and it comes in various forms. All appraisals are a form of *judgment* or *evaluation* about the world, and particularly about the implications of some particular state of the world (or of the self) for the self. When I walk down the street and the grocery bag that I am carrying splits open, spilling its contents onto the sidewalk, I very quickly experience feelings of shock, dismay, embarrassment, and perhaps anger. This complex cascade of emotion is interwoven

³The number varies from three to ten, and researchers often add or subtract emotions from their own lists. See Solomon (2002, p. 123) for an argument that this variety in identifying a canonical number can be a sign of a healthy science and not a weakness.

⁴In basic emotion theory work, one rarely sees an analysis of how a moderately complicated emotional experience like the one associated with lost groceries would proceed. Ellsworth (2013, p. 127) has shrewdly asked “Does Hamlet *ever* feel one of the six basic emotions?”

and shaped by an equally complex set of appraisals, some very fast and automatic, others more controlled and reflective.⁵ In terms of the kinds of content these appraisals have, Ellsworth (2013, p. 125) lists “novelty, valence, certainty, goal conduciveness, agency, and control” as the dimensions most commonly proposed. Thus, upon my noticing the ripping of the grocery bag, the initial reflexive response might be one of novelty (it is not expected) quickly followed by valence (negative). These two responses combine to produce the reaction of shock listed earlier. The list of emotions following shock are shaped by a series of more complex appraisals involving how I view myself, how others view me, and a range of goals in addition to getting my groceries home (e.g. appearing competent). Thus, appraisal theories see emotional experience as being built up, or constructed, out of an often-complex sequence of appraisal and emotional reaction to that appraisal, which includes subjective feeling, physiological processes, and motivations to action. For this reason, Ellsworth (2013) prefers to avoid the word emotion and uses “emotional experiences” instead.

The basic emotions, then, are not so basic (Ortony & Turner, 1990). Some appraisal theorists see these emotional experiences as grouped into clear themes (Lazarus, 1991), while others see the variation as more open ended where appraisal structures can get us “to the right neighborhood, but not the street address” of the experience (Ellsworth, 2013). This understanding of emotion is compatible with the cottage industry of research on “moral emotions” (reviewed in Section 8.2) but does not make the assumptions about a limited set of basic emotions or universality. And it allows much more room for cognition and culture to play a central part in the emotional experience.

8.1.3 Cultural Emotion Theories

It is just as certain that there are cultural differences in emotion as that there are cultural universals. For example, in one Buddhist village in Thailand, it is a moral failure to show emotional involvement when something bad happens: it is better to *tham jai*, to practice a calm heart. But in a nearby Christian village, emotional involvement even to the point of anger is morally appropriate (Cassaniti, 2014). There are numerous examples of these sorts of differences both within national boundaries and across them (see Keltner, Oatley, & Jenkins, 2014, chapter 3; Shweder et al., 2008 for a review). One can also look for variation across time. In Western cultures, religious conceptions made melancholy a preferred emotion in medieval times, while the reformation’s emphasis on a more merciful god had the effect of making smiling a more acceptable, and even pious, public face (Stearns, 2008). Corbin (1986) traces the change in the sense of disgust in France as it helped to establish a cultural divide between the washed and the unwashed. Thus, emotion is linked not only to culture but also to power.

These cultural and historical differences can be quite complex and subtle (Matsumoto et al., 2008). The work on culture and emotion has mostly been interested in documenting these subtleties and in showing how they are related to

⁵See Chapter 7 for a description of “two-process models” that underlie these different kinds of cognitive appraisals.

important themes such as power and morality (Shweder, 2012, 2014a, 2014b). The approach suggests there is likely an infinite number of emotional experiences that a culture or language can highlight and use – that they have structure and are related in thematic groups, although the same emotions do not always occur in the same group in every culture. Finally, the cultural approach encourages us to cast the widest net to understand the role of emotions in morality. It sees emotion as “the entire script” of the experience and its influences including the physiological state (chest pain, goose flesh), the affective experience (panic, emptiness), simple and complex perceptions of the event and the experience (death of a friend, oppression of a minority, victory in struggle), implications for the self (loss, success, empathy), associated social judgment (vice or virtue, sickness or health), and a plan or impetus toward action (withdraw, confess, revenge, comfort). The emotional experience is a continuous experience, of a unit, and though one can look at components,⁶ one cannot adequately consider them in isolation.

8.1.4 Constructionist Models of Emotion

The most recent arrival on the scene of emotion theory are the constructivist approaches (Barrett, 2013; Clore & Ortony, 2013; Cunningham, Dunfield, & Stillman, 2013). These approaches combine elements of the three previous approaches and claim historical precedence to them via one of the founders of psychology, William James (Lindquist, 2013).⁷ They use physiological evidence generated by the basic emotion theories to claim that discrete emotion categories cannot consistently and specifically be localized to distinct brain regions (Lindquist et al., 2012). Their central claim is that *we experience emotions as though they are basic but they are in fact constructed* on the fly by conceptual acts that (1) match the situation with an appropriate, available emotion category and (2) initiate or enact “the embodied experience of the emotion” (Lebois et al., 2020).

Barrett (2013) recalls Darwin’s (1859) emphasis on variability as a central engine of evolution and begins the argument for constructivism by emphasizing the immense variation in appraisal, physiological signature, and expression with any category of emotion, using this variation as a central principle to deconstruct emotion as “not a thing but a category of instances” (Barrett, 2017, p. 16). This emphasis on categorization as the central engine in emotion production that directs bodily expression, situational meaning, action tendencies, etc. allows constructivism to embrace all the complexity of the cultural approaches. One primary distinguishing characteristic between appraisal and constructivist approaches is the amount of expected variation within any particular category, with constructivist approaches allowing for close to

⁶Shweder et al. (2008, p. 415 ff) list eight components, and show their conceptual dependency with a long comparison of *anger/lung lang* in American and Tibetan respondents. The most important component in this instance is “normative social appraisals,” or morality. See Scherer and Peper (2001) for another attempt at a comprehensive listing of components and processes of emotional experience.

⁷One can claim even further historical reach if one recognizes Magda Arnold’s (1960) arguably constructivist approach as grounded in Thomas Aquinas’s medieval scholastic understanding of *habitus* (Cornelius, 2006).

infinite flexibility. Lebois et al. (2020) have demonstrated that people can learn very specifically situated emotions (in this case, *fear of physical harm vs. social rejection* and *anger at physical harm vs. social rejection*). Functional magnetic resonance imaging (fMRI) scans showed the neural activity of the two types of fear or the two types of anger overlapping as little as 11.5%. This kind of emotional learning can underlie tremendous flexibility in emotion while still connecting it to the neural substrate. And it suggests processes by which one might achieve very specific emotional expertise in a domain.

8.1.5 Preliminary Conclusions About Emotions and Morality

Emotions are not simply “feelings” but are instead a complex of: (1) *cognitive conceptualization* (that arouses and shapes the emotion); (2) *directed arousal* (that provides the “feel” of the emotion and signals its urgency); and (3) *action tendency* (that links action to the reaction of the emotion) (see Oatley, Keltner, & Jenkins, 2006 for an overview). What does this picture of emotions tell us about the relations of moral emotion and reason to moral action?

First, there are complex interrelationships between emotion and reason in moral action, rather than a simple opposition.⁸ The appraisal, cultural, and constructivist approaches all see emotion as in part *constituted by reason*. Moral emotions contain, and express the importance of, propositional knowledge. One is angry because an injustice, an important injustice, has occurred. When we learn the injustice was a mistake, or unintentional, our anger is modified. Given this intimate relationship between reason and emotion, there should be no surprise that moral reasoning can inform moral emotion and that moral emotion informs moral reasoning.

Second, the function of emotion in moral action is to “prioritize thoughts, goals, and actions” (Keltner, Oatley et al., 2014, p. 237) in a world where there is rarely enough information to make a complete judgment based on reason or learning alone (Oatley & Johnson-Laird, 2011; Simon, 1967). Thus, rather than being a separate system that biases the primary functions of rational moral judgment, emotion is a basic part of our moral apparatus. It is not an interloper or an outside bias. *If we were not emotional, we would not be able to be fully moral.*

Third, there is a great deal of complexity hidden in our emotional and moral lives that is obscured by the simplistic opposition of reason and emotion. An emotional experience can vary over time and be transmuted, for instance, from envy to anger or sadness (R. Smith & Kim, 2007). Thus, there is likely no canonical list of discrete moral emotions, but rather a complex palette of emotional experience that varies over time and across cultures (Barrett, 2013; Ellsworth, 2013). It still might make sense to study “anger” as though it was a thing, but we should do so knowing how variable the experience is and how porous its boundaries are (if there *are* any boundaries). In addition to not being bounded in basic units, emotions are also not completely bounded within a person. Emotions are not merely personal, but are shared with

⁸See Chapter 2 for the evidence for massive interconnectivity between emotion and cognition at the physiological level. See also Wong (2015) on the convergence between this conclusion of interconnectivity and the philosophical approaches of Mencius and Confucius.

others and this sharing modifies how we and how those others feel (Goldenberg et al., 2015, 2020; E. R. Smith & Mackie, 2008; Wagner et al., 2015).

Fourth and finally, emotions need *not* be the door to relativism. Since they contain propositional knowledge (appraisals) they can be subject to consistency criteria. Even the basic experience of an emotion can feel incorrect (Givon et al., 2022). Consistency criteria can have their effect through reappraisal during the process of the emotional experience (Did they really mean that?) or in the social sharing of emotion (Are you sure they meant that?).⁹ It involves asking the question: Given the relevant circumstances, what is the appropriate emotion? And thus, what appropriate judgment or action proceeds from that? Keane (2015, p. 194) provides rich examples of the social process of rational/emotional argument in several reform movements based on ethical commitments. For example, in the feminist consciousness-raising movement, participants often discovered that, because of unacknowledged injustice, “they had been ‘angry’ for years without realizing it.” Since they are shaped by appraisals, emotional experiences can hang together in predictable patterns that allow for personal examination (Ellsworth, 2013) and philosophical and cultural critique (Gibbard, 2006).

8.2 The Moral Emotions

Some emotions (e.g. anger, compassion) are more closely connected to issues of morality than are others (e.g. sadness, joy). In this section, we will review those emotions and look at how they are woven into our moral lives.

First, however, a word must be said about how one identifies an emotion as belonging to the moral domain. Again, our approach is to be as inclusive as possible. And thus, almost every emotion, from angst to schadenfreude, might be found to have moral implications. But the list here is limited in part by the available research – though the list is growing, only a good handful of emotions have enough work on them to merit inclusion for now. Haidt (2003b, p. 854), one of the pioneers in research on moral emotion, identifies the moral emotions as those meeting two criteria: an emotion is moral to the extent that (1) it is about others (is triggered by the actions or reactions of others, like righteous anger), and to the extent that (2) it produces pro-social motivations.

Even here, though, things get difficult. (1) demarcates emotions as social. But there are, for instance, moral emotions such as gratitude or awe that lead to striving after religious goals (Emmons, 2005, 2008) or to self-transcendence (Koltko-Rivera, 2006) and that are thought of by many religious traditions as deeply moral but not prototypically social.¹⁰ Thus, the elicitors and motivations of a moral emotion may not always be social in nature. (2) marks moral emotions as providing motivation for pro-social behavior. This too is complicated, given that pro-social behavior

⁹But see Chapter 5 for the many kinds of consistency criteria.

¹⁰See Cuttack (1997) and Pratt (1928) for examples of this moral imperative in, respectively, Sufi and Buddhist approaches. Our thanks to David Wulff for providing these examples.

can include altruistic punishment that involves harming others for the sake of reinforcing a moral value.¹¹ Thus, disgust, anger, contempt, shame, pride, and a host of other moral emotions can lead to drastically immoral ends (aided and abetted by reason), resulting in moral collapse (Hart, 2005; Hart & Matsuba, 2007). So both criteria to demarcate moral emotions have difficulties. But they provide a place to begin; as long as we are aware of the limitations.

To organize this review, we will use the divisions into four “emotion families,” suggested by Haidt (2003b): *other-condemning*, *other-suffering*, *self-conscious*, and *other-praising*.¹² These families are more like overlapping sets, with some emotions falling equally into two families (e.g. for this reason we have moved guilt from Haidt’s *self-conscious* family to the *other-suffering* family – it has aspects of both). One can imagine more complicated moral emotions like envy (R. Smith & Kim, 2007) that could fall into three families (*self-conscious*, *other-praising*, and *other-condemning*). Or emotions such as the Thai emotion of *tham jai*, keeping a calm heart, that do not seem to fall easily into any family (Cassaniti, 2014). And one emotion can have different aspects depending on the context. Anger, for instance, plays a central role in close relationships, often contributing to relationship repair (Fitness, 2009, 2015) but shows a different face to threatening out-groups (Fiske, 2015).

The confusion that a close look at the families metaphor produces is a clue that we might not be looking at emotions correctly when we think they are primarily a reaction to one thing. They can be, and often are, a reaction to many things (see the grocery bag example in Section 8.1.1). But even deeper, they are less a reaction and more a *relation* to states of the world and to ourselves. So, anger is not a reaction to an injustice, it is about our relation to an injustice.¹³ The same with compassion, it is about our relation to a person. And as a relation, we have some freedom in shaping it. Even so the families have the advantage of suggesting similarities among groups of emotions, and of suggesting additional investigation when one finds complexities.

In our review, we will follow two aspects of each moral emotion: the emotion as information and as motivation. We know from appraisal research that appraisals are propositional in nature (Wu et al., 2021). For example, “that is unjust” can contain information about relevance based on my goals, intentionality of the actor, amount of harm, etc. Emotions also have a motivational component, or what some theorists have called an “action tendency” (Ellsworth, 2013). They both ready the body for a particular kind of action and orient the individual toward that action. It is this interaction between information and motivation that makes moral emotion a central mover to moral action. Pay attention to this interaction as we review each moral emotion.

8.2.1 Other-Condemning Moral Emotions

These are emotions directed toward the condemnation or rejection of others. They share a focus on negative evaluations of the actions or character of others and an

¹¹See Chapter 1 for more on this concept.

¹²Note that the titles of the categories already include information and motivation as aspects of the emotions.

¹³See Chapter 5 for the variety of ways that this relation can be constituted.

action tendency of rejection or aggression. They often also seem to combine with each other to produce hybrids like moral outrage (Salerno & Peter-Hagene, 2013).

8.2.1.1 Disgust Disgust is an emotion associated with the rejection of unpleasant things. The human facial expression associated with core disgust is the same one displayed when tasting very bitter compounds (Herz, 2011). Even those blind from birth make the same expression of disgust as sighted people. There appear to be three “varieties” of disgust: (1) *core disgust*, related to contact with rotted food, disease, or open wounds; (2) *sexual disgust*, related to rejecting sexual acts; and (3) *moral disgust*, related to rejecting moral transgressions. These three types share some characteristics (e.g. aspects of facial expression) but differ in terms of what provokes them and in the shape of their motivation (Tybur, Lieberman, & Griskevicius, 2009). As with much work in moral emotions, these distinctions are under dispute, with some seeing moral disgust as a mere metaphorical use and others claiming that the types are actually different emotions (see Cameron et al., 2015; Yoder, 2016 for a review). For our purposes, it seems clear that there are at least “loose correspondences” (Cameron et al., 2015) in the form of moral rejection called moral disgust. It is a rejection that is more motivating and less malleable than simple avoidance based on prudence (e.g. disease is disgusting vs. it is prudent to avoid disease). Moral disgust has been shown to have the contamination pattern of core disgust (e.g. people will not want to wear the clothes worn by a murderer, even if they have been cleaned) (Rozin, Haidt, & McCauley, 2008). The evolutionary or cultural origins of these correspondences is, of course, under dispute. But as in many of the controversies in moral psychology, the answer likely lies in some combination of the two perspectives.

We have already mentioned the difficulty of maintaining clear distinctions among emotions. In the moral domain, disgust often overlaps with anger in the rejection of moral harm. Salerno and Peter-Hagene (2013) have documented how disgust and anger combine in moral outrage against particularly heinous offenses. Moral disgust also has well-documented overlap with other emotions, such as contempt (Cameron et al., 2015; Hutcherson & Gross, 2011).

In summary, the information associated with the moral aspect of disgust is that the thing or person is not merely negative but is to be condemned and rejected because it violates, or is associated with the violation of, some moral value. The action tendency is rejection and avoidance that is highly motivating and much less flexible than simple rejection based on prudence. The malleability of disgust, in particular its relation to other moral emotions, is an important open question.

8.2.1.2 Anger There is a voluminous literature on anger, what produces it, how it feels, how it affects the person, and what kind of behavior and goals it produces (Berkowitz & Harmon-Jones, 2004). But the psychological focus on anger as a moral emotion is more recent.¹⁴ We mentioned earlier how it connects and blends with disgust, in its role of reacting to moral violation. Much older work on anger connects it with the appraisal of blocking a goal (Dollard et al., 1939). And this appraisal can

¹⁴Aristotle, of course, saw anger as a moral, social, and indeed political emotion (Sokolon, 2006).

blend into moral anger when we see intention, particularly unjustified intention. We can even find moral anger when we “see” intention in an inanimate object like a car that will not start (Waytz, Cacioppo, & Epley, 2010). So, the information or judgment about morality that moral anger encodes is about moral violation, but more specifically, unjustified moral violation. Scherer (1997) presents a large cross-cultural study in which reported episodes of anger were elicited primarily by appraisals of unfairness and immorality, and consistently more than mere goal frustration.

The *extent* of the moral violation is also associated with the amount of anger and with the associated amount of preferred punishment for the offender (Carlsmith, Darley, & Robinson, 2002). Perpetrators and victims often disagree about the extent of a moral violation and thus about the appropriate response, and this disagreement can escalate into angry exchanges (Baumeister, Stillwell, & Wotman, 1990).

One might think from the comments so far that the primary function of anger is to motivate the victim to punish the offender as a means of self-defense (Darley, Carlsmith, & Robinson, 2000). And indeed, this is one of the motivations of moral anger. But anger can also motivate *altruistic punishment*, defined as the punishment of a moral violator when you yourself are not the victim (Balliet & Van Lange, 2013; Boyd et al., 2003). Thus, one can be angry *for* another person and angry *at* the person who harmed them, with these processes being sometimes separable (Hechler & Kessler, 2018). Things get rapidly more complex when one considers anger in the context of close relationships, where anger and love coexist. In this context, one might say that the motivation is assertion of one’s interest even if one’s interest is the good of the other (Fitness, 2009, 2015). Its complexity is also apparent in long-term social injustice, where despite progress much remains to be angry about (Lorde, 1981). Finally, anger is flexible in its motivational profile as one changes the frame in which the action is interpreted. In forgiveness, one reduces the motivation to punish the offender, but perhaps not one’s bad feelings toward them (Exline et al., 2003). Prayer for the offender can reduce anger, perhaps through cognitive reappraisal (Bremner, Koole, & Bushman, 2011). Anger seems to be more flexible in responding to this kind of reappraisal than other moral emotions are, like disgust (Russell & Giner-Sorolla, 2010), though it appears that disgust and anger can combine in moral outrage (Salerno & Peter-Hagene, 2013).

8.2.1.3 Contempt Like the moral aspects of anger and disgust, moral contempt is about moral rejection, but its scope and shape differ. Moral contempt is about morally loaded incompetence or unfitness. Contempt is a major driver in prejudice and stereotyping (Fiske, 2015) and is (with disgust) the primary emotion shown toward disliked groups perceived as lower in status or less competent. Perceptions of dominance or status are very strongly correlated with judgments of competence, across many cultures (Fiske, 2015), and contempt is the social emotion reserved for those who threaten status in this way. Others call this a community dimension and emphasize the perceived violation of community standards. This alternative dimension involves such things as failure in “duty, role-obligation, respect for authority, loyalty, group honor, interdependence, and the preservation of the community” (Rozin et al., 1999).

We do not know enough about contempt to be able to easily outline the appraisals that lead to it (Cameron et al., 2015). It seems likely that contempt is a complex

emotion associated with perceiving the other as having bad or threatening intentions *and* being low in status or incompetent.¹⁵ This would include violators of duty to the group (e.g. traitors), or those who through willful incompetence within the group threaten it (bumblers), or with out-group or outcast individuals who are viewed as both having bad intent and being less competent. This list allows one to see how contempt can easily be associated with anger for those who endanger us.

The motivational aspect of contempt is similar to disgust: to “mark those who ... represent a threat and avoid them” (Hutcherson & Gross, 2011, p. 720) perhaps by excluding or driving them away. It seems likely that the perception of a failure at intended action is required for the contempt to take on a moral tone (Hutcherson & Gross, 2011) and thus, unlike disgust, contempt seems to be a thoroughly social emotion.

8.2.2 Other-Suffering Moral Emotions

These moral emotions are focused on the suffering of others. But they both offer a stake for the self in the suffering of others, they connect the concerns of the self to the other. They are, then, more complicated than merely being outward focused, but they are focused on the other in a way that motivates the self to action to repair damage or to help or comfort the other.

8.2.2.1 Guilt We feel guilty when we perceive that we have committed a moral violation (Haidt, 2003b, p. 861). More specifically, a moral violation that harms someone else who is important to us (e.g. in a close relationship, group, or community). One often does not feel guilt when harming those who deserve it (see anger, contempt, and disgust, discussed earlier). Modern soldiers, for instance, are well trained in avoiding guilt for the harm they cause to the enemy – though the fog of war and the humanity of the enemy make this complicated (Grossman, 1996).

An important aspect of the appraisal that leads to a feeling of guilt (as opposed to shame) is that it focuses more on the act itself and its victim and less on the self who did the act.¹⁶ Its appraisal has an informational focus outward to the other. With this focus on the other who is harmed, guilt’s motivational influence leads to apology, to restoration of relationship, and to avoidance of future harm (Baumeister, Stillwell, & Heatherton, 1994). Thus, guilt seems to be an emotion that draws people together after harm doing.¹⁷

An important complication in this story is that self-report studies of guilt almost always turn up some instances of non-social moral violations causing guilt. Haidt

¹⁵ See Fiske (2015) for a short review of these two dimensions, their implications in prejudice and stereotyping, and their associations with moral emotions.

¹⁶ Tangney et al. (2007, p. 358) discuss group-based guilt. Because group-based guilt is based in identity, this complicates things. There is also some evidence that the way one measures guilt can tap neurotic vs. pro-social guilt (Tignor & Colvin, 2016).

¹⁷ There is evidence that in some shame-based cultures, pure guilt is a rare and seldom discussed emotion. In these cultures, shame takes the role of reintegrating people into the group and encouraging them to address the harm they have done and reform their behavior (Boiger et al., 2013; Haidt, 2003b).

(2003b, p. 861) speculates that this may be the simple mislabeling of shame as guilt, citing for instance of breaking one's diet. Rather than treating this as a problem of measurement, it would be more productive to investigate the phenomenon itself. It seems possible that in cultures and sub-cultures that value the ethical dimension of divinity or sacredness (Shweder et al., 2008), one can fail at a goal of maintaining purity and, concentrating on the act of failure, feel guilt and distance oneself from one's action.¹⁸ Whether this requires a "person" to whom one reconciles is unclear. In this case, avoiding strict definitions of morality (e.g. as necessarily social) may help us discover important and complex aspects of moral functioning in some cultures.

8.2.2.2 Compassion In their textbook on the emotions, Keltner, Oatley et al. (2014, p. 219) call compassion the "master social emotion" that helps to "form stronger social bonds of all kinds." In this context, they cite the popular historian of religion Karen Armstrong (2006) and her claim that compassion is central to all the major religious traditions. As with each emotion we review here, there is some controversy surrounding its form and centrality (Goetz, Keltner, & Simon-Thomas, 2010).¹⁹ But there appears to be a growing consensus that compassion does indeed serve a central role in pro-social emotions (Stellar et al., 2011).

Compassion is the feeling of concern for another, accompanied by a desire to enhance that person's welfare.²⁰ But more specifically, it is the result of a cascade of appraisals about the eliciting situation (Goetz et al., 2010; Stellar et al., 2011).²¹ First, it focuses on the suffering of another person who is relevant to the self. This "relevant to the self" aspect is important to understand the unique motivating power of compassion. But it also helps to understand the wide variation in who is considered eligible for compassion. The Oliners (1988) have, for instance, documented how an extensive vision of who the in-group is helped to shape the actions of rescuers in the

¹⁸Buber (1958) makes a distinction between guilt, a social moral emotion, and sin, an emotion associated with the self having violated the sacred. They can be conceptually distinguished, but can they be empirically distinguished? It might be useful to ask what informational and motivational components are associated with sin. Is the emotional experience of sin more like guilt, leading the person to reparation, or more like shame, leading to negative self-focus and withdrawal? Here is some fascinating territory for interaction between theology and psychology.

¹⁹For example, see Chapter 1 for discussion on how a "selfish" evolutionary process might produce compassion as a central emotion in a social species.

²⁰Compassion need not always be directed outward. Neff (2003a, 2003b), drawing on Buddhist approaches, has defined self-compassion as having three aspects: self-kindness, a sense of common humanity, and mindfulness when evaluating the self and its weaknesses. Self-compassion may be a better predictor of self-worth and well-being than self-esteem scales (Barnard & Curry, 2011b; Neff & Vonk, 2009) and is associated with resilience in service workers, such as pastors (Barnard & Curry, 2011a).

²¹Note that this series of appraisals are remarkably like those in the early models of helping (Latane & Darley, 1970). This similarity makes sense given that taking action to help is the primary motivational aspect of compassion.

Holocaust.²² Second, there is an appraisal of deservingness of compassion. Compassion is often restricted when the person harmed has in some way contributed to or “deserved” that harm (Rudolph et al., 2004). Third, there is an appraisal of the ability one has to help. Because there is such a strong motivation to action in compassion, it is often reduced to the extent that one is helpless to intervene. This suppression of experiencing compassion when one cannot help puts stress on seeing oneself as a moral person (Cameron, Harris, & Payne, 2016; Cameron & Payne, 2012) and can be relieved only by some trade-off of one’s moral commitments or one’s perception of the victim. Thus, even though (or perhaps because) compassion is the “master social emotion” it contains a complicated and contradictory set of appraisals and information about the world.

Compassion “derives its strength from the emotional stake it offers the self in the other’s welfare” (Preston & de Waal, 2002, p. 279). The combination of appraisals in compassion makes it uniquely motivating: that the person is indeed suffering, that they are deserving of compassion, and that one can help. These connect the self to the other in a powerful way. The strength is mirrored in the physical expression characteristic of compassion, the head orientation toward the victim and a forward lean (Goetz et al., 2010).

Compassion gives us an example of the complexity of the appraisal process and its effects on the shape and strength of a moral emotion. Researchers have documented similar complexity in the anger–disgust–contempt triad. It is a good sign that psychological research in moral emotions is now entering the domain of human complexity in relations. And as always, there is much more to be done.

8.2.3 Self-Conscious Emotions

This family of emotions has the self as the primary focus. In embarrassment, pride, and shame, one is aware of good or bad aspects of the self. All other moral emotions (indeed, all emotions) implicate the self too, in that they mark how some state of the world is relevant for the self. But this family is also focused on the self in a self-reflexive fashion.²³ One self-consciously reflects on the self. Still, even these inward turned moral emotions are social: embarrassment requires a kind of “social exposure,” pride and shame can both be deeply implicated in social comparison.

8.2.3.1 Embarrassment At first glance, embarrassment might not seem relevant to the moral domain, and some authors (Lewis, 2008) divide embarrassment into types, one of which is mere “social exposure” without any negative implications. But even mere social exposure can carry the threat of losing status in a group, being singled out for no good reason, or even being singled out with praise that one worries is more than one deserves (or will be tolerated). Thus, embarrassment is a self-focused emotion that is about one’s belongingness and obligations to the

²² See Chapters 4 and 5 for greater detail about how these individual differences in compassion play out in moral action.

²³ It can easily become more complex. In subsequent rounds of appraisal, any other emotion can become caught in self-reflection.

group.²⁴ It is pro-social in that it motivates reintegration into the group.²⁵ But even this motivation for reintegration is embedded in the self's moral expectations about the self and its place in the social world.

The information that embarrassment carries for the individual experiencing it is that there is a threat of loss of status in one's group or society or of an uncomfortable status differential. For instance, in the classic Japanese novel *Tale of Genji* (Murasaki, [ca. 1000]) embarrassment is a central emotion that occurs sometimes because of improper behavior, but sometimes merely because one person is in the presence of a social superior (Keltner, Oatley et al., 2014, p. 78). Some theorists see these as two different forms of embarrassment with the improper behavior form appearing as a less intense version of shame (Lewis, 2008). Other authors point out the crucial cultural embeddedness of embarrassment and shame and suggest that, in cultures in which the self is deeply interdependent with others, shame and embarrassment "merge together into a single emotion of immense moral importance" (Haidt, 2003b, p. 859).

The primary moral motivation associated with embarrassment is to signal submission to others (Keltner & Buswell, 1997) or to be conciliatory and conforming (Tangney, Stuewig et al., 2007). It is like shame in that some of its facial expressions (e.g. averted eyes) express anxiety about evaluation, but it is differentiated from shame in that the desire to reintegrate with the group is stronger than the desire to hide (Keltner & Buswell, 1997).

8.2.3.2 Pride Pride generally comes from an appraisal that something good is associated with the self. It is a deeply self-conscious emotion, and thus requires a sense of self. For this reason, even though it emerges in early childhood, it is only anchored in a detailed and inward sense of self in adolescence (Hart & Matsuba, 2007). It is often thought of as occurring in domains that are not necessarily moral, such as social recognition and occupational achievement. Here we will concentrate on pride that is associated with moral being or doing. Moral pride is, then, the emotional reward when there is an achievement in the moral domain as suggested by Hart and Matsuba (2007).

But there is more complexity here. Recent work distinguishes between two distinct types of pride, each of which can be about moral action: *hubristic* and *authentic* (Dickens & Robins, 2022; Tracy & Robins, 2007). Hubristic moral pride is based in a judgment about the global self and its moral achievements, and it is often (but not necessarily) comparative. Authentic pride is occasioned by specific moral achievement and its connection to the self and its values (Hart & Matsuba, 2007).²⁶

²⁴See Chapter 7 for moral foundations (Graham et al., 2011) that involve loyalty to the group. Embarrassment may be seen as a recognition of a threat to this value.

²⁵And it interestingly shares neural substrate with other pro-social emotions such as compassion (Moll et al., 2007).

²⁶In a large cross-sectional international sample (Orth, Robins, & Soto, 2010), authentic pride increases from adolescence into old age, while hubristic pride decreases from adolescence to adulthood, reaching a minimum around age sixty-five, and then begins to increase. This suggests a kind of moral maturation in this domain.

The information in the emotion of pride differs depending upon the type. For hubristic pride, it is a sense of supremacy of the global self over others (Carver & Johnson, 2010; Tangney, Stuewig et al., 2007). It is usually comparative (Tangney, Stuewig et al., 2007) and related to social dominance (Carver & Johnson, 2010). Hubristic moral pride would come from thinking, for instance, that one (and perhaps one's group) was more authentic, caring, and principled than others. It is about moral superiority that is explicitly comparative. It is a pride of "being" (Lewis, 2008).

In contrast, authentic moral pride is a pride that stems from doing good. It results from the perception that one is responsible for a specific positive moral outcome, or has resisted a particular temptation, or acted in regard to particular obligations (Malti & Ongley, 2014).

The motivational complex associated with hubristic pride is (not surprisingly) negative and even anti-social. Hubristic pride is associated with impulsivity, anger, aggression, seeking public recognition, narrow focus on financial success, and social dominance (Carver & Johnson, 2010; Liu et al., 2016). In the business domain, it is associated with decisions to pursue unwise mergers and acquisitions that destroy value rather than create it (Bodolica & Spraggon, 2011). Because of the impulsivity and self-centeredness, hubristic pride makes the maintenance of relationships (business or otherwise) difficult. It also serves a central emotional role in prejudice and intergroup bias (Fiske, 2015) and leads people to identify only with powerful others (Oveis, Horberg, & Keltner, 2010). Authentic moral pride is associated with self-control, adaptive achievement, and goal engagement (Carver & Johnson, 2010) and is a primary motive for "voluntary planned, sustained moral action" that reflects the self and its values (Hart & Matsuba, 2007, p. 117).

If the distinctions among these types of pride seem too strongly drawn and simplistic, it is in part because they are: the literature is still young and developing. But there are interesting complexities. Bodolica (2011) recognizes, for instance, the dynamism that hubristic pride can give an organizational leader, and Carver and Johnson (2010) list several aversive emotional experiences that tend not to correlate with either type of pride (e.g. anxiety/fear). There is also some suggestion that gratitude plays an essential role in how pride is expressed. Organizational leaders who experience pride in achievement tend to treat their employees well only if they feel grateful for employees' contributions to that success (Michie, 2009). Finally, we should note that, though hubristic pride is implicated in prejudice toward out-groups (Fiske, 2015), there is very little published on hubristic *moral* pride, the pride that comes from thinking oneself morally superior.

8.2.3.3 Shame Shame is a moral emotion that is complicated by cross-cultural difference. In most Western cultures, shame comes from the appraisal that something is wrong or inadequate with the self. But in most non-Western cultures, shame is part of a complex of embarrassment, shyness, modesty, and social fear, and is crucial to life embedded in social groups (Haidt, 2003b; Kitayama & Park, 2007). For instance, shame is a somewhat negative but culturally condoned emotion in Japan. Japanese schools have times set aside for *hansei*, to encourage children to take critical self-reflection in the service of self-improvement (Boiger et al., 2013). Shame in the United States, however, is to be avoided because it threatens positive self-regard

(Haidt, 2003b). Thus, one can expect that the information embedded in shame and the motivational profile of shame will be quite different in Western vs. non-Western cultures.²⁷

In Western cultures, shame involves negative evaluation of the self, and often includes imagery of how one appears to others in this light (Tangney, Stuewig et al., 2007). It might be occasioned by something the person has done, but it can also be elicited by appearing inadequate, flawed, or defective (Haidt, 2003b). Thus, while shame is an inward-focused moral emotion, it is in part about how this inward self appears to the outer social world. The inward focus is made clear by the motivations associated with shame: hiding, distress about the self, anger at others who make one look bad (Stuewig et al., 2010; Tangney, Stuewig et al., 2007). It is this withdrawing, defensively aggressive (to self and other) reaction that makes shame problematic in Western cultures. It leads one away from social engagement. Criminals who focus on shame during their imprisonment are *more* likely to commit additional crimes and return to jail compared with those who focus on guilt (Tangney, Mashek, & Stuewig, 2007).²⁸ The withdrawal motivation associated with shame is part of what discourages victims of sex abuse from reporting incidences and the focus on negative self-evaluation makes shame a central driver in depression long after sexual abuse has taken place (Lewis, 2008).²⁹

In non-Western cultures, shame is occasioned when one has violated social expectations (e.g. by asking for or needing help) or is in danger of violating social expectations (e.g. by simply being in the presence of someone higher ranking) (Boiger et al., 2013; Haidt, 2003b). In such cultures, there is no clear distinction between moral norms and social conventions (Shweder et al., 1997). Though still negative, the experience of shame in these cultures is less intense than in Western cultures (Boiger et al., 2013) and it serves as a motivation to reintegrate oneself into the group. It is pro-social, less about the core self and more about behavior that can be remedied. We are just beginning to tease apart these complexities.

²⁷The differences are not simply dichotomous. One can find them across time and within shared cultures. For instance, as Western cultures have moved away from communal and toward individualist societies, they have also de-emphasized shame. There are large differences within Japan in individualism (Kitayama et al., 2006), and one can place some cultures in the middle on the scale of independence–interdependence, thus implicating differences in the associated emotions – e.g. Kitayama et al. (2006) find that Britain and Germany occupy a space in between the United States and Japan on that scale.

²⁸The relationship of shame to criminal reform is still not well understood. The research program of Tangney and colleagues cited here suggests shame is problematic for reform. But a central theoretical perspective in restorative justice approaches is “reintegrative shaming” that focusses on making public disapproval of the crime explicit, with the goal of reintegrating the criminal (Harris, 2006). Perhaps the emphasis on “reintegrative” helps explain this disconnect in the literatures.

²⁹In the Western world, *shaming* might be understood as an act of moral violence, making a public show of focusing on the low worth of the shamed person. One can think of this as a punishment motive, using the emotion of shame to hurt the target and make the target of shaming feel bad.

8.2.4 Other-Praising Moral Emotions

This emotion family recognizes positive characteristics in others, and responds with a range of reactions that strengthen relationships (Haidt, 2003a). Though they are outward focused, it is unclear whether the “other” need be human, or even personal. And they may also be as closely related to an inward-focused relation to the self, as seen in the relationship between gratitude, awe, and humility (Kruse et al., 2014; Piff et al., 2015).

8.2.4.1 Elevation and Awe Elevation is the emotional response to witnessing acts of virtue or moral beauty (Algoe & Haidt, 2009). Its characteristic feeling is that of outward-directed expansion, perhaps even uplifting and inspiring (Haidt, 2000). This description of the feeling suggests a motivational state, which is outward directed. Those experiencing elevation often want to praise the other person, share news of the good deed, or even emulate it. Those experiencing elevation are motivated to become like the other, at least in action.³⁰

The subjective experience of elevation is different from simply feeling good or feeling happy, and evidence suggests that this difference leads to clear increases in pro-social activity on the part of those experiencing the emotion (Schnall, Roper, & Fessler, 2010). For example, in a study of Italian business leaders, it has been shown that leaders’ interpersonal fairness and self-sacrifice leads to the emotion of elevation in employees, and that the experience of this emotion is a clear influence on employees’ organizational commitment and citizenship (Vianello, Galliani, & Haidt, 2010). It seems reasonable to expect such an emotional effect to be limited to the domain in which it occurs (e.g. the workplace), but there is no research that documents this.

The experience of elevation, and its effects, is also influenced by one’s readiness to be open to it. People higher in moral identity, for instance, are more likely to experience elevation when they see morally praiseworthy action, and they are more likely to emulate that action (Aquino, McFerran, & Laven, 2011).³¹ The link with moral identity and the subjective experience of elevation as a positive emotion together suggest that the motivation elevation produces is not one of overcoming some deficit or lack but instead one of approaching something desirable.³² However, it seems possible that the desirable here need not always be the moral. It may be that group socialization toward prejudice occurs in part because of a desire to emulate others (see Gatto & Dambrun, 2012 for an example involving police officers). We know very little about how flexible the experience of elevation is regarding its (im)moral goal.

Finally, a note on the related emotion of awe, though it is not clear that this emotion fits the category of “other-praising” (Shiota, Keltner, & Mossman, 2007). Awe stems from an appraisal of vastness that is not graspable, possibly accompanied by

³⁰The emotion that leads to the desire to emulate the other is treated by Zagzebski (2013, 2017) as a central component in a theory of virtue ethics.

³¹This is again a place where the idea of a single focus of an emotion misleads us. It is more the *relation* of the self to the awe-inspiring than simply the outward focus. One finds this theme in much religious poetry and writing (e.g. Rumi, Meister Eckhardt) and the relational aspect of awe is reflected in the current empirical work.

³²See Chapters 5 and 9 for the various ways one might be motivated to moral action.

threat, exceptional beauty, virtue, and the supernatural (Keltner & Haidt, 2003). Because it is associated with a situation that “overwhelms current cognitive categories” awe has a complicated relationship with the information it provides. In some studies the experience of awe increases belief in supernatural agency, partly by motivating a desire for cognitive closure (Valdesolo & Graham, 2014). But in other studies it has been shown that those without a need for cognitive closure are more open to experiencing awe (Shiota et al., 2007). But there is agreement that, rather than being simply an enjoyable experience, awe increases pro-social behavior, that is, it has an action tendency. How it does so is still unclear – perhaps by reframing one’s view of the self (Keltner, Kogan et al., 2014). And this reframing suggests again the dual nature of focus of these emotions: both other-praising and inward focused.

8.2.4.2 Gratitude One may experience gratitude when one receives some benefit from another, particularly if it is clear that the benefactor is acting out of a correct awareness of the needs of the recipient (Algoe & Haidt, 2009; Bono, Emmons, & McCullough, 2012). Like elevation, gratitude’s elicitors and its effects are more specific than simple happiness (Kruse et al., 2014). Experiencing gratitude motivates one to want to return the favor, strengthen the relationship with the benefactor, or even to pass benefit on to others (Bono et al., 2012; Emmons & Kneezel, 2005; McCullough et al., 2001).

One *may* experience gratitude, but there are a variety of influences that lead people away from this experience, even if they do receive benefits from others. These include perceptions of victimhood, envy or resentment, or an emphasis on materialistic values (see Bono et al., 2012 for a review). There is some evidence that insecure relationship attachment styles can lead to lower levels of gratitude (Dwiwardani et al., 2014). It makes sense that skepticism about relationships (being wary of or blaming others or being cynical about others’ motivations) should also reduce gratitude. Still, one should make the distinction between *non*-gratitude and *ingratitude*. Non-gratitude may simply be the result of forgetting or paying attention to other things, but ingratitude has a more aggressive edge, and may even be motivated by a desire to punish or hurt the other (Emmons, 2007, p. 141).

In addition to motivating people to draw closer to each other, gratitude can have other social effects. For instance, gratitude and humility are linked in a virtuous cycle over time, in which one’s gratitude influences one’s view of the self as dependent on others (Kruse et al., 2014), but this cycle is likely to work only when one’s dependence is not seen as a burden (Bono et al., 2012).³³ In Section 8.2.4.1, we reviewed work that showed that leaders’ gratitude to their workers was an important influence in inspiring those workers to emulate their leaders (Vianello et al., 2010). Thus, gratitude calls forth elevation and emulation.

³³See Weidman, Cheng, and Tracy (2018) for an explanation of the type of humility that might best fit this virtuous cycle (appreciative vs. self-abasing). One can also find this distinction in Thomas Merton’s (2009) commentary on the Rule of Benedict (Benedict of Nursia, 550/1980).

8.3 Learning Moral Emotions

From our review of moral emotions so far, one can see that having passions does not mean we are passive in our emotional experiences. We are constantly engaged in the appraisal that constructs them. As infants, we come with emotion built in, but much of the rest of life involves our grappling with our emotions, learning to guide, control, produce, and integrate them into our plans and actions. We do this because when our emotion regulation goes out of control, our emotion can indeed be overwhelming. But we also do so in part because when they are working well our emotions make our life rewarding. As we move from childhood to adolescence, we become aware of the complexities of relationships and situations, which results in increasingly complex moral emotions and cognition. We learn to recognize emotion in ourselves, to distinguish one emotion from another, to recognize mixed emotions, and to evaluate our emotions for appropriateness. There is, then, a developmental progression in the moral emotions.³⁴ And there is an identifiable way to track emotional learning at the neural level and to gauge its variability (Lebois et al., 2020).

For example, consider the “Happy victimizer” phenomenon. As they develop and learn moral rules, children can say that some particular action (like stealing a toy) is bad. But early in this process they are also willing to say that a child who does the wrong action will feel happy, because the child gets his or her desire (the toy). It is only later, as they begin to come to terms with how emotions in others work, that they recognize that emotions can be linked both to getting what you want and to feeling bad if you break a rule. The extent to which one makes the connection between feeling bad and having done something wrong can predict delinquency in adolescence (Krettenauer, 2013). Children may not have developed the skills associated with theory of mind to anticipate other and self-reactions to a transgression. But even adults may not see morality as centrally relevant to the self or as relevant to some particular situation (Malti & Ongley, 2014, p. 167).

Even in adulthood we can learn emotions. A meditation-based compassion training program, in which one practices feeling compassion for a range of people, can increase altruistic action and change emotional experience (Weng et al., 2013). We can also learn to be more grateful (Bono et al., 2012). Consciousness-raising groups (e.g. in the early feminist movement) achieve their purpose in part through training emotional responses, such as learning when anger is appropriate (Keane, 2015). Even though there is clear evidence that our nonconscious intuition and emotion can drive moral judgment, specific attempts to reappraise our moral emotional responses can alter emotion and overturn these judgments (Feinberg et al., 2012). We can rethink (and refeel) our moral emotional reactions. Most forms of psychotherapy have at least some emotional component, and insight-based therapies that concentrate on emotion recognition and reappraisal produce longer lasting effects in clients than more cognitive approaches (Keltner, Oatley et al., 2014, p. 373ff.). Rivers et al. (2013) have developed a successful training program in classrooms to help increase what they

³⁴See Keltner, Oatley et al. (2014, chapter 8) for a broad but concise review of the development of emotions.

call “emotional intelligence” – the set of skills that allow pro-social emotions to support good classroom environments.

So, of course we can rethink our moral emotional reactions. We are not passive in the face of our passions. We should acknowledge, however, that we can rethink our emotions in order to be better killers in battle (Grossman, 1996) and to reduce our distress when we harm others (Bandura, 1999) as well as to be more compassionate. But given the right circumstances, we can learn, and unlearn, our moral emotions. This in turn means that we are not at the mercy of our emotions, not the passive victims of them, not always led astray by them. In fact, it means that one of the ways we are moral is by being emotional about those values to which we are passionately committed.

8.4 Discussion

8.4.1 Conclusion

1. *Emotions are not simply biases, they are evaluations that can be appropriate or inappropriate.*

Are emotions always biasing to moral judgment and action? In short, no. Both emotion and reason *can* bias moral judgment, but it is not inevitable. In fact, the complex interactions among emotion and judgment are the reason we can be moral. Common to almost all definitions of emotion is the mixture of information and judgment, physiological reaction, subjective feel, and motivational profile.³⁵ This complex interweaving of emotion and reflection, of appraisal, feeling, and action tendency, is what makes moral life work in the individual psyche, and in social interaction. We are moved to action by emotion, and our emotions are informed by judgment. This rapid switching among thoughts and feelings, interrupting, conversing, and blending,³⁶ gives us some idea of the constant interweaving of reason and emotion in our moral life.

Rather than moral emotions biasing evaluations, moral emotions *incorporate* a significant aspect of evaluation, and are deeply entwined in moral judgment and action. Indeed, emotions can only *be* biasing because they contain cognition in the form of information, judgment, and motivation. Nor is the strength of the

³⁵ See Keltner, Oatley et al. (2014, p. 9) for at least nine different definitions of emotion. Definitions are less important than exploring the related processes within their contexts. It is more important to understand how any particular process works and the variations in how it works. So, in this chapter we have cast a wider net, loosened from definitional constraints, both in terms of emotion and the moral.

³⁶ In Shakespeare’s *Winter’s Tale*, King Leontes narrates his thoughts and feelings as he comes to believe (falsely) that his wife has cheated on him: “Then’tis very credent, Thou may’st co-join with something, and thou do’st, (And that beyond commission) and I find it, (And that to the infection of my brains, And hardening of my brows).” For those not facile with Shakespearian English, here is a transliteration showing the rapid switching among the thoughts: “It is then very believable, you may have co-joined with something [King Polixenes], And you do! (and you do it outside my authority!) and I have found it!, and it causes me to be crazy, and to become angry.”

emotion a good guide to whether it will bias moral evaluation. Even very strong emotional reactions (e.g. righteous anger, moved by compassion) can be morally beneficial. The gentler passions also can be biasing (e.g. the self-pity that allows me to distance myself from those I harm).

2. *We are not as passive in experiencing our moral emotions as we often believe.*

Emotions, and perhaps especially moral emotions, are not experiences we have, they are performances we do. Emotional responses can be learned and can be modified by reappraisal. Consciousness-raising groups are settings where people learn what are considered appropriate moral emotional reactions. Compassion meditation has clear effects on shared emotion. Given a particular moral ecology and appropriate experience, perhaps we can become expert feelers.

3. *We should include emotion in models of reflective equilibrium in moral evaluation.*

A reasonable goal here might be called reflective equilibrium among emotions and rationality (Daniels, 2018; Narvaez, 2010; Rawls, 1971/1999). Since emotions can be tested for their appropriateness to a situation, and reasons and reasoning can interact with this testing process, one can learn the skills of moral discernment by involving both emotion and reason. Emotion-free rationality without a search for this equilibrium should be just as suspect as emotion-laden judgment.³⁷ But because emotions are social things, we need to add a social dimension to reflective equilibrium. One can see this in a little discussed aspect of Haidt's (2001) model in which the moral emotions and reactions of one's peers influence, often not consciously, one's own moral emotions and reactions. The Japanese practice of using shame in *hansei* provides an opportunity for schoolchildren to learn appropriate moral emotions. Adding this social aspect suggests *social-emotional reflective equilibrium* as a process and goal for mature, or expert, moral action.

4. *Moral emotion should be integrated into the teaching of ethics.*

Finally, since moral emotions are integral in moral action, what implication should this have for how we teach ethics? Perhaps we should reverse a long standing trend in ethics education that presumes that solving ethical puzzles is the best training (Pincoffs, 1971) and begin to make ethics emotional again. To do so would involve (un)consciously training emotion – learning when to trust it, how to guide it, how to argue with (and for) it, and when and how to question it. Perhaps practicing appropriate emotional responses and their reappraisal through social-emotional reflective equilibrium should be at the center of ethics training. Maybe it already is at the center of ethics training but unacknowledged and not well integrated.

8.4.2 Application

The case of moral courage in the face of pressure to overturn the 2020 presidential election is replete with moral evaluation, where one finds both imbalance and attempts at balance in social-emotional reflective equilibrium. Donald Trump himself is a paradigmatic example of someone driven by emotion, intuition, and an often-self-serving

³⁷In Chapter 7, we look at how reason itself can go awry to bias moral judgment. In Chapter 9, we look at how one might cultivate this social-emotional reflective equilibrium.

desire who uses reason only as the servant of his goals. Aaron Van Langevelde was in a role on the Michigan Board of State Canvassers that required him to give reasons for his actions, and so he is unlikely to offer emotional explanations. But if one looks closely at the reasons he gives, the urgency and difficulty of his position is easy to see. He voted to certify the election results in Michigan to avoid “constitutional chaos and the loss of our integrity” (all quotes are from his speech at Cardozo Law School in Boston where he received an award for professional courage, Bidgood, 2021). He explained that he was being asked by Trump and others to “disregard [his] oath of office,” “abandon [his] administrative role,” “ignore a clear legal duty,” and “take power [he] didn’t have.” Emotion-loaded words like chaos, loss, disregard, abandon, and ignore all point to the emotional pull of his professional identity as a public servant but also as a lawyer for the Republican Party in Michigan. These words invoke emotions like fear, indignation, contempt, and scorn but also loyalty and devotion. That the words he chooses are emotion-laden is not a mistake or bias in his evaluation of the situation. In software design language, the emotions are not bugs but features. They motivate him in the social–emotional reflective equilibrium to overcome the intense pressure he describes to be loyal to Trump. He mentions an outpouring of social pressure to vote no or to not attend the certification, even to the point that “my family and I were in danger.” But despite this pressure, he felt the “rule of law is a worthy cause.” This is an evaluation, driven by a balance of emotion and reason, socially validated by the award he received. He was surely not doing it to get the award (he was its first recipient); but the award was designed to offer public praise to shore up a shaking moral ecology.

8.4.3 Open Questions

1. *How can we treat the wide confusion within and between emotion categories as an asset rather than a liability?*

It seems sufficiently established that we will need to live with the wide plasticity within and between emotion categories like anger and awe. But how can we turn this into a theoretical asset rather than a confession of failure? We suggest it is by embracing the complexity of emotions from the neurological to the cultural level and finding solvable puzzles within this complexity. One way to do this is to study the process of establishing social–emotional reflective equilibrium. Another is by following the developmental progression of emotional complexity over time within particular domains (e.g. How do doctors and therapists learn to properly emote in their fields?). Several other possibilities are listed for this in Section 8.4.3.2.

2. *How do we establish social-emotional reflective equilibrium?*

How might we do research on how and when people use a process like socio-emotional reflective equilibrium? Perhaps by looking at expert feelers. We might study those who are in jobs that require an integration of moral emotion and rationality in the service of getting things done. In short, by studying those who do emotional work (Grandey, 2000), work that involves the regulation of emotion as a part of the job. We can find them in war (Grossman, 1996), hospice nursing (Hill, 2005), counseling (Fowers, 2012; Fowers & Winakur, 2014), and even in food service (Grandey, 2000). Hill (2005), for instance, provides a moving example of the care taken by a team of nurses with a very difficult hospice patient. Much

of their effort focused on maintaining compassion while making rational plans to deal with their own emotional reactions to constant provocation. They were regularly consulting both emotion and reason in the service of moral goals, compassionate care for the patient and concern for their own emotional welfare.³⁸ Or perhaps we might look at morally exemplary actors, who have been studied in a wide range of professions and volunteer positions (Hart, Murzyn, & Archibald, 2013). A common finding among those who do work on moral exemplars is that these actors do not generally do lengthy rational deliberation in the course of their moral careers (Huff & Furchert, 2016). These, too, may be expert at the integration of reason and emotion in the service of moral goals. We might also look at training programs in monastic formation in Christian (Casey, 2005), Hindu (Rambachan, 2006), and other traditions with long experience in training for personal and spiritual growth. And we might look to clinical psychology for systematic work and long experience in personal growth (Weinstein, Przybylski, & Ryan, 2013). Thus, there are places one can find relevant insight, but there is not much work central to current moral psychology.

3. *How do moral emotions play a role in prioritizing moral thoughts, goals and action?* The complex world of moral action is one in which there is rarely enough information to make a complete judgment based on reason or learning alone. Emotions move us both to thought and to action. They do not belong to a separate system that only interferes with the primary functions of rational moral judgment. They are instead a basic part of our moral apparatus and functioning. They are a feature, not a bug. If we were not emotional, we would not be able to be fully moral.

Much work in moral emotions has so far consisted in establishing the importance of moral emotion in moral thought and action. We need to begin analyzing the role they play as they interact with thoughts, goals, and actions, influencing them and being in turn influenced. How, for instance, do emotion and cognitive processing interact in consciousness-raising groups, *hansei*, compassion meditation, intergroup conflict, and religious experience?

8.5 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section. Please look to each section on a specific emotion for relevant references.

- Scherer and Peper (2001). "Psychological theories of emotion and neurological research." A neuroscience-based review of emotions that provides one view of the aspects of an emotion.
- Shweder et al. (2008). The cultural psychology of emotions: Ancient and renewed." A review of cultural aspects of emotion that provides a comprehensive list of aspects of emotion.

³⁸Note that this team approach is social, emotional, and rational all at once, and it seeks to maintain a balance based upon a moral goal: compassionate professional care.

- Keltner, Oatley et al. (2014). *Understanding Emotions* (3rd ed.). A textbook-length, comprehensive review of work on emotions.
- Ellsworth (2013). "Appraisal theory: Old and new questions." One of the primary appraisal theorists offers a critique and directions for work in the area.
- Cameron et al. (2015). "A constructionist review of morality and emotions." A review of emotion literature that finds no evidence for specific links between moral content (e.g. harm, purity) and discrete emotions (e.g. anger, disgust).
- Barrett (2021). "This is how your brain makes your mind." An approachable and clear overview of how emotions are constructed "on the fly."

References

- Algoe, S. B., & Haidt, J. (2009). Witnessing excellence in action: the "other-praising" emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology*, 4(2), 105–127. doi:10.1080/17439760802650519
- Aquino, K., McFerran, B., & Laven, M. (2011). Moral identity and the experience of moral elevation in response to acts of uncommon goodness. *Journal of Personality and Social Psychology*, 100(4), 703–718. doi:10.1037/a0022540
- Armstrong, K. (2006). *The Great Transformation: The Beginning of Our Religious Traditions*. New York: Knopf.
- Arnold, M. B. (1960). *Emotion and Personality*. New York: Columbia University Press.
- Balliet, D., & Van Lange, P. A. M. (2013). Trust, punishment, and cooperation across 18 societies: A meta-analysis. *Perspectives on Psychological Science*, 8(4), 363–379. doi:10.1177/1745691613488533
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193.
- Barnard, L. K., & Curry, J. F. (2011a). The relationship of clergy burnout to self-compassion and other personality dimensions. *Pastoral Psychology*, 61(2), 149–163. doi:10.1007/s11089-011-0377-0
- Barnard, L. K., & Curry, J. F. (2011b). Self-compassion: Conceptualizations, correlates, & interventions. *Review of General Psychology*, 15(4), 289–303. doi:10.1037/a0025754
- Barrett, L. F. (2013). Psychological construction: The Darwinian approach to the science of emotion. *Emotion Review*, 5(4), 379–389. doi:10.1177/1754073913489753
- Barrett, L. F. (2017). *How Emotions are Made: The Secret Life of the Brain*. Boston: Mariner Books.
- Barrett, L. F. (2021). This is how your brain makes your mind. *MIT Technology Review* (August). Retrieved from <https://www.technologyreview.com/2021/08/25/1031432/what-is-mind-brain-body-connection/>
- Barrett, L. F., Adolphs, R., Marsella, S. et al. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. doi:10.1177/1529100619832930
- Baumeister, R. F., Stillwell, A., & Wotman, S. R. (1990). Victim and perpetrator accounts of interpersonal conflict: Autobiographical narratives about anger. *Journal of Personality and Social Psychology*, 59(5), 994–1005.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: an interpersonal approach. *Psychological Bulletin*, 115(2), 243–267.
- Benedict of Nursia. (550/1980). *The Rule of Benedict* (T. Fry, Trans.). Collegeville, MN: Liturgical Press.

- Berkowitz, L., & Harmon-Jones, E. (2004). Toward an understanding of the determinants of anger. *Emotion*, 4(2), 107–130. doi:10.1037/1528-3542.4.2.107
- Bidgood, J. (2021, July 5). Aaron van Langevelde's speech about the 2020 election: "We were asked to take power we didn't have." *Boston Globe*.
- Bodolica, V., & Spraggon, M. (2011). Behavioral governance and self-conscious emotions: Unveiling governance implications of authentic and hubristic pride. *Journal of Business Ethics*, 100(3), 535–550. doi:10.1007/s10551-010-0695-7
- Boiger, M., Mesquita, B., Uchida, Y. et al. (2013). Condoned or condemned: The situational affordance of anger and shame in the United States and Japan. *Personality and Social Psychology Bulletin* 39(4), 540–553. doi:10.1177/0146167213478201
- Bono, G., Emmons, R. A., & McCullough, M. E. (2012). Gratitude in practice and the practice of gratitude. In P. A. Linley & S. Joseph (Eds.), *Positive Psychology in Practice* (pp. 464–481). New York: Wiley.
- Boyd, R., Gintis, H., Bowles, S. et al. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6), 3531–3535. doi:10.1073/pnas.0630443100
- Bremner, R. H., Koole, S. L., & Bushman, B. J. (2011). "Pray for those who mistreat you": Effects of prayer on anger and aggression. *Personality and Social Psychology Bulletin*, 37(6), 830–837. doi:10.1177/0146167211402215
- Buber, M. (1958). *Schuld und Schuldgefühle*. Heidelberg: Heidelberg Press.
- Cameron, C. D., Harris, L., T., & Payne, B. K. (2016). The emotional cost of humanity: Anticipated exhaustion motivates dehumanization of stigmatized targets. *Social Psychological and Personality Science*, 7(2), 105–112.
- Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review*, 19(4), 371–394. doi:10.1177/1088868314566683
- Cameron, C. D., & Payne, B. K. (2012). The cost of callousness: regulating compassion influences the moral self-concept. *Psychological Science*, 23(3), 225–229. doi:10.1177/0956797611430334
- Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, 83(2), 284–299.
- Carver, C. S., & Johnson, S. L. (2010). Authentic and hubristic pride: Differential relations to aspects of goal regulation, affect, and self-control. *Journal of Research in Personality*, 44(6), 698–703. doi:10.1016/j.jrp.2010.09.004
- Casey, M. (2005). Models of monastic formation. In M. Casey (Ed.), *An Unexciting Life: Reflections of Benedictine Spirituality* (pp. 365–405). Petersham, MA: St. Bede's Publications.
- Cassaniti, J. L. (2014). Moralizing emotion: A breakdown in Thailand. *Anthropological Theory*, 14(3), 280–300.
- Clare, G. L., & Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emotion Review*, 5(4), 335–343. doi:10.1177/1754073913489751
- Cohen, J. D. (2005). The vulcanization of the human brain: A neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives*, 19(4), 3–24.
- Corbin, A. (1986). *The Foul and the Fragrant: Odor and the French Imagination*. Cambridge, MA: Harvard University Press.
- Cornelius, R. R. (2006). Magda Arnold's Thomistic theory of emotion, the self-ideal, and the moral dimension of appraisal. *Cognition & Emotion*, 20(7), 976–1000. doi:10.1080/02699930600616411
- Cunningham, W. A., Dunfield, K. A., & Stillman, P. E. (2013). Emotional states from affective dynamics. *Emotion Review*, 5(4), 344–355. doi:10.1177/1754073913489749

- Cuttick, W. C. (1997). Rumi and the Mawlawiyyah. In S. H. Nasr (Ed.), *Islamic Spirituality: Manifestations* (pp. 105–126). New York: Crossroad Publishing.
- Daniels, N. (2018). Reflective Equilibrium. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2018/entries/reflective-equilibrium/>
- Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, 24(6), 659–683.
- Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (1st ed.). London: John Murray.
- Dickens, L. R., & Robins, R. W. (2022). Pride: A meta-analytic project. *Emotion*, 22(5), 1071–1087. doi:<http://dx.doi.org/10.1037/emo0000905>
- Dollard, J., Miller, N. E., Doob, L. W. et al. (1939). *Frustration and Aggression*. New Haven, CT: Yale University Press.
- Dwiwardani, C., Hill, P. C., Bollinger, R. A. et al. (2014). virtues develop from a secure base – attachment and resilience as predictors of humility, gratitude, and forgiveness. *Journal of Psychology and Theology*, 42(1), 83–90.
- Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 45–60). New York: Wiley.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3(4), 364–370. doi:10.1177/1754073911410740
- Ellsworth, P. C. (2013). Appraisal theory: Old and new questions. *Emotion Review*, 5(2), 125–131. doi:10.1177/1754073912463617
- Emmons, R. A. (2005). Sacred purposes: What they are and why they matter. *Psychology of Religion Newsletter*, 20(1), 1–7.
- Emmons, R. A. (2007). *Thanks! How the New Science of Gratitude Can Make You Happier*. Boston: Houghton Mifflin Harcourt.
- Emmons, R. A. (2008). Gods and goals: Spiritual striving as purposeful action. *Journal für Psychologie*, 16(3).
- Emmons, R. A., & Kneezel, T. T. (2005). Giving thanks: Spiritual and religious correlates of gratitude. *Journal of Psychology and Christianity*, 24(2), 140–148.
- Exline, J. J., Worthington, E. L., Hill, P. et al. (2003). Forgiveness and justice: A research agenda for social and personality psychology. *Personality and Social Psychology Review*, 7(4), 337–348. doi:10.1207/s15327957pspr0704_06
- Feinberg, M., Willer, R., Antonenko, O. et al. (2012). Liberating reason from the passions: overriding intuitionist moral judgments through emotion reappraisal. *Psychological Science*, 23(7), 788–795. doi:10.1177/0956797611434747
- Fiske, S. T. (2015). Intergroup biases: a focus on stereotype content. *Current Opinion in Behavioral Sciences*, 3(June), 45–50. doi:10.1016/j.cobeha.2015.01.010
- Fitness, J. (2009). Anger in relationships. In H. T. Reis & S. Sprecher (Eds.), *Encyclopedia of Human Relationships* (pp. 94–97). Thousand Oaks, CA: Sage.
- Fitness, J. (2015). Emotions in relationships. In M. Mikulincer & P. R. Shaver (Eds.), *APA Handbook of Personality and Social Psychology: Vol. 3. Interpersonal Relations* (Vol. 3, pp. 297–318). Washington, DC: American Psychological Association.
- Fowers, B. J. (2012). Placing virtue and the human good in psychology. *Journal of Theoretical and Philosophical Psychology*, 32(1), 1–9. doi:10.1037/a0025819
- Fowers, B. J., & Winakur, E. (2014). Key virtues of the psychotherapist. In S. Van Hooft (Ed.), *The Handbook of Virtue Ethics* (pp. 386–396). New York: Routledge.
- Freud, S. S. (1905/1962). *Three Essays on the Theory of Sexuality* (J. Strachey, Trans.). New York: Basic Books.
- Gatto, J., & Dambrun, M. (2012). Authoritarianism, social dominance, and prejudice among junior police officers. *Social Psychology*, 43(2), 61–66. doi:10.1027/1864-9335/a000081
- Gibbard, A. (2006). Moral feelings and moral concepts. *Oxford Studies in Metaethics*, 1, 195–215.

- Givon, E., Udelsman-Danieli, G., Almagor, O. et al. (2022). Can feelings “feel” wrong? Similarities between counter-normative emotion reports and perceptual errors. *Psychological Science*, 33(6), 948–956. doi:10.1177/09567976211063915
- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin*, 136(3), 351–374. doi:10.1037/a0018807
- Goldenberg, A., Garcia, D., Halperin, E. et al. (2020). Collective emotions. *Current Directions in Psychological Science*, 29(2), 154–160. doi:10.1177/0963721420901574
- Goldenberg, A., Halperin, E., van Zomeren, M. et al. (2015). The process model of group-based emotion: Integrating intergroup emotion and emotion regulation perspectives. *Personality and Social Psychology Review*, 20(2), 118–141. doi:10.1177/1088868315581263
- Graham, J., Nosek, B. A., Haidt, J. et al. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101, 366–385.
- Grandey, A. A. (2000). Emotion regulation in the workplace: A new way to conceptualize emotional labor. *Journal of Occupational Health Psychology*, 5(1), 95–110.
- Grossman, D. (1996). *On Killing: The Psychological Cost of Learning to Kill in War and Society*. New York: Back Bay Books.
- Haidt, J. (2000). The positive emotion of elevation. *Prevention and Treatment*, 3(1), 1–5. doi:10.1037/1522-3736.3.1.33c
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J. (2003a). Elevation and the positive psychology of morality. In C. L. M. Keyes & J. Haidt (Eds.), *Flourishing: Positive Psychology and the Life Well Lived* (pp. 275–289). Washington, DC: American Psychological Association.
- Haidt, J. (2003b). The moral emotions. In R. J. Davidson, Scherer, K. R., & Goldsmith, H. H. (Ed.), *Handbook of Affective Sciences* (pp. 852–870). Oxford: Oxford University Press.
- Harris, N. (2006). Reintegrative shaming, shame, and criminal justice. *Journal of Social Issues*, 62(2), 327–346.
- Hart, D. (2005). The development of moral identity. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (pp. 166–196). Lincoln, NE: University of Nebraska Press.
- Hart, D., & Matsuba, M. K. (2007). The development of pride and moral life. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The Self-Conscious Emotions: Theory and Research* (pp. 114–133). New York: Guilford Press.
- Hart, D., Murzyn, T., & Archibald, L. (2013). Informative and inspirational contributions of exemplar studies. In M. K. Matsuba, P. E. King, & K. C. Bronk (Eds.), *Exemplar Methods and Research: Strategies for Investigation: New Directions for Child and Adolescent Development* (pp. 75–84). New York: Wiley.
- Hechler, S., & Kessler, T. (2018). On the difference between moral outrage and empathic anger: Anger about wrongful deeds or harmful consequences. *Journal of Experimental Social Psychology*, 76, 270–282. doi:10.1016/j.jesp.2018.03.005
- Herz, R. S. (2011). PROP taste sensitivity is related to visceral but not moral disgust. *Chemosensory Perception*, 4(3), 72–79. doi:10.1007/s12078-011-9089-1
- Hill, R. (2005). End-of-life care for the patient with borderline personality disorder. *Journal of Hospice and Palliative Nursing*, 7(3), 150–161.
- Huff, C., & Furchert, A. (2016). Embracing the complexity of moral action. In D. Bertini (Ed.), *Moral Heteronomy. History, Proposals, Arguments, a Special Issue of Dialektika*.
- Hume, D. (1739/1985). *A Treatise of Human Nature*. New York: Penguin.
- Hutcherson, C., A., & Gross, J. J. (2011). The moral emotions: A social-functionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, 100(4), 719–737.
- Keane, W. (2015). *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton University Press.

- Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin*, 122(3), 250–270.
- Keltner, D., & Haidt, J. (2003). Approaching Awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion*, 17(2), 297–314. doi:10.1080/02699930244000318
- Keltner, D., Kogan, A., Piff, P. K. et al. (2014). The sociocultural appraisals, values, and emotions (SAVE) framework of prosociality: Core processes from gene to meme. *Annual Review of Psychology*, 65, 425–460. doi:10.1146/annurev-psych-010213-115054
- Keltner, D., Oatley, K., & Jenkins, J. M. (2014). *Understanding Emotions* (3rd ed.). New York: Wiley-Blackwell.
- Kierkegaard, S. (1849/2004). *The Sickness Unto Death* (A. Hannay, Trans.). New York: Penguin.
- Kitayama, S., Ishii, K., Imada, T. et al. (2006). Voluntary settlement and the spirit of independence: Evidence from Japan's "northern frontier." *Journal of Personality and Social Psychology*, 91(3), 369–384. doi:http://dx.doi.org/10.1037/0022-3514.91.3.369
- Kitayama, S., & Park, H. (2007). Cultural shaping of self emotion and well being: How does it work? *Social and Personality Psychology Compass*, 1(1), 202–222.
- Koltko-Rivera, M. E. (2006). Rediscovering the later version of Maslow's hierarchy of needs: Self-transcendence and opportunities for theory, research, and unification. *Review of General Psychology*, 10(4), 302–317. doi:10.1037/1089-2680.10.4.302
- Krettenauer, T., Campbell, S., & Hertz S. (2013). Moral emotions and the development of the moral self in childhood. *European Journal of Developmental Psychology*, 10(2), 159–173. doi:10.1080/17405629.2012.762750
- Kruse, E., Chancellor, J., Ruberton, P. M. et al. (2014). An upward spiral between gratitude and humility. *Social Psychological and Personality Science*, 5(7), 805–814. doi:10.1177/1948550614534700
- Latane, B., & Darley, J. M. (1970). *The Unresponsive Bystander: Why Doesn't He Help?* Upper Saddle River, NJ: Prentice Hall.
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8), 819–834.
- Lebois, L. A. M., Wilson-Mendenhall, C. D., Simmons, W. K., Barrett, L. F., & Barsalou, L. W. (2020). Learning situated emotions. *Neuropsychologia*, 145(August), 106637. doi:10.1016/j.neuropsychologia.2018.01.008
- Lewis, M. (2008). Self-conscious emotions: Embarrassment, pride, shame, guilt. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 742–756). New York: Guilford Press.
- Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: a modern psychological constructionist model. *Emotion Review*, 5(4), 356–368. doi:10.1177/1754073913489750
- Lindquist, K. A., Wager, T. D., Kober, H. et al. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(3), 121–143. doi:10.1017/S0140525X11000446
- Liu, C., Yuan, L., Chen, C. et al. (2016). Authentic and hubristic pride as assessed by self, friends, and strangers. *Social Psychological and Personality Science*, 7(7), 690–696. doi:10.1177/1948550616649240
- Lorde, A. (1981). The uses of anger. *Women's Studies Quarterly*, 9(3), 7–10.
- Malti, T., & Ongley, S. F. (2014). The development of moral emotions and moral reasoning. In M. K. J. Smetana (Ed.), *Handbook of Moral Development* (2nd ed., pp. 163–183). New York: Psychology Press.
- Matsumoto, D., Seung Hee, Y., & Fontaine, J. (2008). Mapping Expressive differences around the world: the relationship between emotional display rules and individualism versus collectivism. *Journal of Cross-Cultural Psychology*, 39(1), 55–74. doi:10.1177/0022022107311854

- McCullough, M. E., Kilpatrick, S. D., Emmons, R. A. et al. (2001). Is gratitude a moral affect? *Psychological Bulletin*, 127(2), 249–266.
- Merton, T. (2009). *The Rule of Saint Benedict: Initiation into the Monastic Tradition 4*. Collegeville, MN: Liturgical Press.
- Michie, S. (2009). How positive emotions influence the prosocial behaviors of organizational leaders. *Journal of Leadership & Organizational Studies*, 15(4), 393–403.
- Moll, J., de Oliveira-Souza, R., Garrido, G. J. et al. (2007). The self as a moral agent: linking the neural bases of social agency and moral sensitivity. *Social Neuroscience*, 2(3–4), 336–352. doi:10.1080/17470910701392024
- Murasaki, S. [ca. 1000]. *The Tale of Genji* (K. Suyematz, Trans.). Cambridge, ON: In Parentheses Publications.
- Narvaez, D. (2010). Moral Complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Neff, K. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2(2), 85–101. doi:10.1080/15298860390129863
- Neff, K. D. (2003b). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250. doi:10.1080/15298860390209035
- Neff, K. D., & Vonk, R. (2009). Self-compassion versus global self-esteem: Two different ways of relating to oneself. *Journal of Personality*, 77(1), 23–50. doi:10.1111/j.1467-6494.2008.00537.x
- Nesse, R. M. (2014). Comment: A general “theory of emotion” is neither necessary nor possible. *Emotion Review*, 6(4), 320–322. doi:10.1177/1754073914534497
- Oatley, K., & Johnson-Laird, P. N. (2011). Basic emotions in social relationships, reasoning, and psychological illnesses. *Emotion Review*, 3(4), 424–433. doi:10.1177/1754073911410738
- Oatley, K., Keltner, D., & Jenkins, J. M. (2006). *Understanding Emotions*. New York: Wiley-Blackwell.
- Oliner, S. P., & Oliner, P. M. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press
- Orth, U., Robins, R. W., & Soto, C. J. (2010). Tracking the trajectory of shame, guilt, and pride across the life span. *Journal of Personality and Social Psychology*, 99(6), 1061–1071. doi:10.1037/a0021342
- Ortony, A., & Turner, T. J. (1990). What’s basic about basic emotions? *Psychological Review*, 97(3), 315 – 331. doi:10.1037/0033-295X.97.3.315
- Oveis, C., Horberg, E. J., & Keltner, D. (2010). Compassion, pride, and social intuitions of self-other similarity. *Journal of Personality and Social Psychology*, 98, 618–630.
- Piff, P. K., Dietze, P., Feinberg, M. et al. (2015). Awe, the small self, and prosocial behavior. *Journal of Personality and Social Psychology*, 108(6), 883–899. doi:http://dx.doi.org/10.1037/pspi0000018
- Pincoffs, E. (1971). Quandry ethics. *Mind & Language*, 80, 552–571.
- Pratt, J. B. (1928). *The Pilgrimage of Buddhism*. New York: Macmillan.
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1), 1–20. doi:10.1017/s0140525x02000018
- Rambachan, A. (2006). *The Advaita Worldview: God, World, and Humanity*. New York: State University of New York Press.
- Rawls, J. (1971/1999). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rivers, S. E., Brackett, M. A., Reyes, M. R. et al. (2013). Improving the social and emotional climate of classrooms: a clustered randomized controlled trial testing the RULER Approach. *Prevention Science*, 14(1), 77–87. doi:10.1007/s1121-012-0305-2
- Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 757–776). New York: Guilford Press.

- Rozin, P., Lowery, L., Imada, S. et al. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76(4), 574–586.
- Rudolph, U., Roesch, S., Greitemeyer, T. et al. (2004). A meta-analytic review of help giving and aggression from an attributional perspective: Contributions to a general theory of motivation. *Cognition & Emotion*, 18(6), 815–848. doi:10.1080/02699930341000248
- Russell, P. S., & Giner-Sorolla, R. (2010). Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science*, 2(4), 360–364. doi:10.1177/1948550610391678
- Salerno, J. M., & Peter-Hagene, L. C. (2013). The interactive effect of anger and disgust on moral outrage and judgments. *Psychological Science*, 24(10), 2069–2078. doi:10.1177/0956797613486988
- Scherer, K. R. (1997). The role of culture in emotion-antecedent appraisal. *Journal of Personality and Social Psychology*, 73(5), 902–922.
- Scherer, K. R., & Peper, M. (2001). Psychological theories of emotion and neurological research. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 5, pp. 17–48). Amsterdam: Elsevier.
- Schnall, S., Roper, J., & Fessler, D. M. T. (2010). Elevation leads to altruistic behavior. *Psychological Science*, 21(3), 315–320. doi:10.1177/0956797609359882
- Shiota, M. N., Keltner, D., & Mossman, A. (2007). The nature of awe: Elicitors, appraisals, and effects on self-concept. *Cognition & Emotion*, 21(5), 944–963. doi:10.1080/02699930600923668
- Shweder, R. A. (2012). Anthropology's disenchantment with the cognitive revolution (1). *Topics in Cognitive Science*, 4(3), 354–361. doi:10.1111/j.1756-8765.2012.01199.x
- Shweder, R. A. (2014a). Let me tell you a story about Hindu temples and runaway trolleys. In D. Sharma (Ed.), *Psychoanalysis, Culture, and Religion: Sudhir Kakar* (pp. 272–282). Oxford: Oxford University Press.
- Shweder, R. A. (2014b). The tower of appraisals: Trying to make sense of the one big thing. *Emotion Review*, 6(4), 322–324.
- Shweder, R. A., Haidt, J., Horton, R. et al. (2008). The cultural psychology of emotions: Ancient and renewed. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 409–427). New York: Guilford Press.
- Shweder, R. A., Much, N. C., Mahapatra, M. et al. (1997). The “big three” of morality (autonomy, community, and divinity) and the “big three” explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and Health* (pp. 119–169). New York: Routledge.
- Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, 74(29–39).
- Smith, E. R., & Mackie, D. M. (2008). Intergroup emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 428–439). New York.
- Smith, R., & Kim, S. H. (2007). Comprehending envy. *Psychological Bulletin*, 133(1), 46–64. doi:10.1037/0033-2909.133.1.46
- Sokolon, M. K. (2006). *Political Emotions: Aristotle and the Symphony of Reason and Emotion*. DeKalb, IL: Northern Illinois University Press.
- Solomon, R. C. (2002). Back to basics: On the very idea of “basic emotions”. *Journal for the Theory of Social Behaviour*, 32(2), 115–144.
- Stearns, P. N. (2008). History of emotions: Issues of change and impact. In M. Lewis, J. Haviland-Jones, M., & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 17–31). New York: Guilford Press.
- Stellar, J. E., Manzo, V. M., Kraus, M. W. et al. (2011). Class and compassion: Socioeconomic factors predict responses to suffering. *Emotion*, 12(3), 449–459. doi:10.1037/a0026508
- Stuewig, J., Tangney, J. P., Heigel, C. et al. (2010). Shaming, blaming, and maiming: Functional links among the moral emotions, externalization of blame, and aggression. *Journal of Research in Personality*, 44(1), 91–102. doi:10.1016/j.jrp.2009.12.005

- Tangney, J. P., Mashek, D., & Stuewig, J. (2007). Working at the social-clinical-community-criminology interface: The George Mason University inmate study. *Journal of Social and Clinical Psychology, 26*(1), 1–21.
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology, 58*, 345–372.
- Tignor, S. M., & Colvin, C. R. (2016). The interpersonal adaptiveness of dispositional guilt and shame: A meta-analytic investigation. *Journal of Personality, 85*(3), 341–363.
- Tracy, J. L., & Randles, D. (2011). Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review, 3*(4), 397–405. doi:10.1177/1754073911410747
- Tracy, J. L., & Robins, R. W. (2007). The psychological structure of pride: a tale of two facets. *Journal of Personality and Social Psychology, 92*(3), 506–525. doi:10.1037/0022-3514.92.3.506
- Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology, 97*(1), 103–122. doi:10.1037/a0015474
- Vaccaro, A., Kaplan, J. T., & Damasio, A. (2020). Bittersweet: The neuroscience of ambivalent affect. *Perspectives on Psychological Science, 15*(5), 1187–1199. doi:10.1177/1745691620927708
- Valdesolo, P., & Graham, J. (2014). Awe, uncertainty, and agency detection. *Psychological Science, 25*(1), 170–178. doi:10.1177/0956797613501884
- Vianello, M., Galliani, E. M., & Haidt, J. (2010). *Elevation at Work: The Effects of Leaders' Moral Excellence*. University of Padova.
- Wagner, U., Galli, L., Schott, B. H. et al. (2015). Beautiful friendship: Social sharing of emotions improves subjective feelings and activates the neural reward circuitry. *Social Cognitive Affective Neuroscience, 10*(6), 801–808. doi:10.1093/scan/nsu121
- Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who sees human?: The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science, 5*(3), 219–232. doi:10.1177/1745691610369336
- Weidman, A. C., Cheng, J. T., & Tracy, J. L. (2018). The psychological structure of humility. *Journal of Personality and Social Psychology, 114*(1), 153–178. doi:10.1037/pspp0000112
- Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2013). The integrative process: New research and future directions. *Current Directions in Psychological Science, 22*(1), 69–74. doi:10.1177/0963721412468001
- Weng, H. Y., Fox, A. S., Shackman, A. J. et al. (2013). Compassion training alters altruism and neural responses to suffering. *Psychological Science, 24*(7), 1171–1180. doi:10.1177/0956797612469537
- Wong, D. B. (2015). Growing virtue: The theory and science of developing compassion from a mencian perspective. In B. Bruya (Ed.), *The Philosophical Challenge from China* (pp. 23–58). Boston, MA: MIT Press.
- Wu, Y., Schulz, L. E., Frank, M. C. et al. (2021). Emotion as information in early social learning. *Current Directions in Psychological Science, 30*(6), 468–475. doi:10.1177/096372142111040779
- Yoder, A. M. (2016). The word disgust may refer to more than one emotion. *Emotion, 16*(3), 301–308. doi:10.1037/emo0000118
- Zagzebski, L. (2013). Moral exemplars in theory and practice. *Theory and Research in Education, 11*(2), 193–206. doi:10.1177/1477878513485177
- Zagzebski, L. (2017). *Exemplarist Moral Theory*. Oxford: Oxford University Press.

Moral Formation: Shaping Moral Action

we are inquiring not in order to know what virtue is, but in order to become good, since otherwise our inquiry would have been of no use.

(Aristotle, *Nicomachean Ethics*, Book II, chapter 2)

In each of the preceding chapters we suggested how the contexts, influences, and processes covered there would support the processes of becoming the kind of person who takes moral action. But we rarely explicitly addressed those long-term processes of becoming that form people into moral actors. For instance, in Chapter 5, we constructed a table (Table 5.1) to unfold the complexities in the criteria we use to evaluate consistency in our behavior. What we did not address in that chapter were the actual processes by which those moral evaluations were made, integrated into one's self, and then supported taking moral action. This chapter is where we take those sometimes isolated pieces and attempt to place them in a larger picture that we call *moral formation*. In order to understand these varying ways of progression or regression in moral action, moral formation must integrate, for each individual actor, the contexts, influences, and processes we have covered in preceding chapters and to which we will refer throughout this chapter.

We borrow the term *formation* from two sources: First, as it appears in the psychological literature on identity formation.¹ And second, from traditional accounts across religions and philosophies as they provide guidance on the formation of the moral self.

Identity formation is that set of processes that help us form our sense of who we are, touching on the question from Chapter 4: "What does a person like me do in a situation like this?" We understand *moral* (identity) formation to be the

¹That literature has until now been primarily concerned with issues of self-discovery and self-creation in adolescence but has expanded its inquiry to emerging adulthood (Lapsley & Hardy, 2017). What these identity formation processes are, and where they lead, are still quite open questions. See Lapsley (2020) for a review of the issues in the field.

continual forming of embodied moral identity, hence that set of processes that helps us form who we are as moral actors, by constant critique and practice. This is a focus on the enduring but mercurial self (moral identity), stitched together by skilled practice and by narrative unity over time. By embodied moral identity we mean to gesture toward the ways that skills and habits and embeddedness in moral ecology in part constitute moral action over time. This formation of moral identity adds the dimension of ongoing practice in concrete contexts to the understanding of the moral self as a cognitive/emotional network. To provide a glimpse of the complexity of these processes, we will look to three traditional accounts of moral formation to help us develop a psychological framework that also addresses moral formation.

In the understanding we develop here, *taking* moral action is central to, and even a part of, moral formation. Thus, moral formation becomes the place where many of the contexts, influences, and processes presented here are integrated. In this framework, we understand moral formation as that set of processes that integrate the diverse influences, contexts, and processes of moral action, that commit the self to moral goals, that evaluate progress toward them, and then make and implement plans, expectations, policies, etc. to shape and support future action.

An example here may help: In the Introduction we presented the case of the Munson family who began by taking in foster children and, through ever-increasing commitments, eventually became social entrepreneurs in their community, founding a large non-profit organization providing support to disabled children. Their moral action over time led to increasing commitments, to the practice and development of cognitive, emotional, and social skills, and to the construction of institutions in the moral ecology. They expanded their circle of moral concern but also limited it with a focus on particular issues and people. Their moral identity, that is, who they are as moral persons and how they act, is bound up with and constituted by all these things.² One can say as much for each of the other cases we present in the Introduction.

One guiding principle of this book is that there are multiple layers, dimensions, or aspects to how we become moral actors.³ Though there is little work that concentrates specifically on *moral* formation (at least in adults),⁴ we have an embarrassment

²The reader should note that these definitions allow for commitment to immoral and amoral goals as well, and do not speak to whether those goals are concrete (support for disabled children in my community) or abstract (compassion).

³In the introduction, we called these different views or approaches *narratives*, to emphasize how they might be interdependent narratives rather than competing theories and to reduce the temptation of reductionism.

⁴Work in character formation is almost exclusively focused on childhood and early adolescence (Lapsley & Hardy, 2017) and work on spiritual formation and spiritual care is fragmented and not well integrated with the psychological literature. For some hints on how these latter two fields might be connected to current work in psychology, see Section 9.1.

of theoretical riches from neighboring research programs in personality and social psychology. However, none of them speak to the unity, richness, and variety of the processes of moral formation that we cover here.⁵ Instead of proposing yet another model, we offer the beginnings of a framework that attempts to grasp the phenomenon of moral formation in its complexity and richness: conscious and nonconscious processes, emotional and cognitive aspects, individual and social levels, habitual and deeply reflective action. These are drawn from the previous chapters with the addition of current psychological research and theory that we consider relevant to underlying processes of moral formation.

In the face of this lack of psychological theory of moral formation, we look in Section 9.1 to several traditional accounts of moral formation as we can find them in: (1) religious and secular approaches to *Bildung*; (2) existential concepts of value/goal appropriation; and (3) traditions of spiritual and monastic formation across the major religions. From these traditions we deduce three shared themes of moral formation to sketch a first landscape of moral formation and to help us gather existing literature, integrate previous chapters, and develop further research questions. With *evaluative-transformative aspects* of moral formation (Section 9.2) we point to those psychological processes that guide how we evaluate, critique, and practice moral commitments and action. Chapters 4, 5, and 6 have already covered many of the processes in this aspect of moral formation. The *normative/self-transcending* (see Section 9.3) moral formation aspects unfold those psychological processes by which we appropriate and integrate normative values into our identity and lives, construct our hierarchies of moral commitments (which in itself has evaluative aspects as we need to evaluate and re-evaluate our value system), and transform and transcend our moral self toward the moral good. Third, with *facilitative/educative* (see Section 9.4) moral formation we explore aspects of moral formation which have to do with education broadly construed (e.g. not just schooling or training), pointing to moral ecologies that foster and facilitate (and impede) moral development and formation. We will look back at moral ecologies and ask how moral ecology shapes and is shaped by moral formation. In the process, we will see that these three aspects do not describe independent types of moral formation but rather aspects of a synthesis of processes that form our moral identity toward moral action.

⁵For early work on choosing goals (moral or not) see Murray (1938), Maslow (1954), and McClelland (1987). More recent work focuses on many aspects, including autonomy (Ryan & Deci, 2003), self-regulation (Carver & Scheier, 2002), executive functioning (Baumeister, Masicampo, & Vohs, 2011), action control (Kuhl & Koole, 2004), goals and plans (Gollwitzer, 2015), goal strivings (Emmons, 1986), life projects (Little, 2008), life tasks (Cantor et al., 1987), prospection (Gilbert & Wilson, 2007; Seligman et al., 2013), future self (van Gelder, Hershfield, & Nordgren, 2013), ideal self (Higgins, Tykocinski, & Vookles, 1990), and an updated evolutionary version of Maslow's model (Ackerman & Bargh, 2010; Kenrick et al., 2010; Peterson & Park, 2010). See also Vaillant (2015) who reports on a study of Western male adult development that by itself spans the decades-long time frame of all these literatures. Narvaez (2010) provides a short overview of the complexity of attempting to integrate these diverse literatures.

9.1 Traditional Accounts of Moral Formation

Since humanity began telling stories, those stories have included themes of moral growth and degeneration.⁶ World and indigenous religions provide guidance for the formation of spiritual and moral identity or self (Narvaez et al., 2019). And, of course, philosophical systems East and West have concerned themselves with how one becomes, and accordingly does, good.

In Sections 9.1.1–9.1.3, we present three traditions of moral formation in suggestive outline and urge the enterprising reader to look more carefully into each of them for clues and critiques of psychological research on moral formation. We hope these will serve as test cases of the intimacy and complexity of the processes and illustrate the need for integrative theory-making in the field.

9.1.1 *Bildung*

There is a German-language narrative tradition of *Bildung* (education as personal formation) that has produced an entire genre of literary narrative, the *Bildungsroman*, the novel of personal formation (Golban, 2018), first clearly exemplified (and self-critiqued) by Goethe's *Wilhelm Meisters Lehrjahre* (1795/2007). This genre is rooted in a long German tradition that refers to *Bildung*, a word for which “there is no economical translation” (Herdt, 2019, p. 6). According to the theological ethicist Jennifer Herdt, the Rhineland mystic Meister Eckhart, writing in the German vernacular, first used the German term *Bild* (picture or image) as a central concept in spiritual/moral formation. Influenced by a theological anthropology that views the person as an image of God, he claims the human task is one of reforming (*überbilden*) and informing (*einbilden*) one's self toward the God-image. The word *Bildung* holds both *Bild* as a noun and *bilden* as a verb, which literally means “the process of shaping” someone or something toward an image, or as the title of our chapter has it, formation. It is a process of forming oneself, and allowing oneself to be formed, after an example (the *Vorbild*). It holds the religious resonances of the Judaic creation story (and Christian theological anthropology) in which humans are created (*gebildet*) in the image (*Bild*) of God. *Bildung* is often described as the educational process of encounter with the world that leads to the development of self, based on an image of the ideal (Stojanov, 2012). But it is more complex than that and has a long history encompassing both variation and contradiction within it.

It is, then, more profound than simple education. At least as articulated by Wilhelm von Humboldt (1792/1960) who transformed higher education in Germany, *Bildung* is the continual attempt of individuals, through experience or interaction with the world, to open up themselves to the world, to relate to that world, to be changed by it, and in particular to become open to the many different perspectives that others have on the world (Stojanov, 2012). This is a goal and vision of personal becoming that emphasizes human emotional and spiritual commitment in encounter with others. It shares things in common with Kohlberg's insistence on wide

⁶See our comment about the moral aspects of the 4,000-year-old Gilgamesh story in Chapter 3.

encounter with the social world (see Chapter 7) but includes a much broader set of values, emotions, and skills within it, making change within the self—identity formation—a central aspect of the process. Finally, it opens the door to moral influence from reading and theater and from that part of our moral ecology that can engage the self from across the centuries of a tradition.

This history is laid out in all its complexity by Herdt (2019). She traces the idea back to classical Greek *paideia* and Roman *humanitas*, an aristocratic project of educating the self for citizenship. Medieval theologians adapted this by appropriating the image of God as the goal for development, whence it passed into the German vernacular as various versions of *bilden* in Meister Eckhart's writings. The tradition explodes in variation and creativity from there, with pietists (Arndt, Boehme), philosophers (Hegel, Herder), secular writers and humanists (Goethe, Schiller), and Pedagogues (Wilhelm von Humboldt) shaping and reshaping it in a dense weave of influence and disagreement.

Herdt (2019) detects a variety of themes of *Bildung*/formation for which we can only give a quick list here: (1) *Bildung* as both gift and task, either religious as in coming from and returning to God, or secular as in emerging from the religion of art/nature; (2) *Bildung* as radical, independent self-construction; (3) *Bildung* as the effort of development into full humanity in loving association with others; (4) *Bildung* as an expansion of the self by encounter with the wide variety of human diversity, conditions, and opinions; (5) *Bildung* as a passive openness to being changed either by God or by the forces immanent in nature.⁷

The richness of the *Bildung* tradition is that it has been seen as a passive and an active process, a religious and a secular endeavor, an individualist and a communitarian education, and a mystical/emotional and an intellectual achievement. But all the variations share the idea of development into a full humanity by intentional encounter over time with something other than the self that produces a powerful transformation of the self.⁸ These variations are better seen as perspectives or tensions within the tradition rather than contradictions. Thus, *Bildung* serves as a model for a moral psychology that itself incorporates tensions of individual autonomy vs. situational pressure, of emotion vs. reason, and of collectivist vs. individualistic culture; And also as a corrective for a moral psychology that has until now avoided grappling with issues of the ultimate.

Herdt also sees these themes re-emerging in recent work by philosophers Martha Nussbaum and Anthony Appiah and their struggle to construct an ethics that recognizes humanity in all its diverse and brilliant vibrancy rather than the flat, colorless language of basic human rights (Appiah, 2001; Nussbaum, 1994). One can also see these themes and tensions in the American version of the liberal arts education and its emphasis on whole person learning, language learning, and study abroad (Hutchins, 1953). We turn now to a tradition that examines in detail the “powerful transformation of the self” on which *Bildung* is based.

⁷Most of these themes appear and are critiqued in *Meister Wilhelm's Lehrjahre* (Goethe, 1795/2007), and this complexity is part of its status as a major work of art.

⁸It must be said that this process is prone to failure, as seen in the incorporation of the German *Bildung* ideal into the National Socialist (Nazi) agenda.

9.1.2 Existential Appropriation

Since ancient times there has been a tradition in philosophy emphasizing Socratic maieutic over the mere acquisition of knowledge and skill. This art of philosophical midwifery focuses on facilitating the birth of truth in the single individual. Such truth is different from objective knowledge or doctrine since it is mediated by an embodied process of existential appropriation,⁹ which refers to a range of internal processes like evaluating and re-evaluating, choosing or struggling to choose, incorporating and embodying what “concerns us ultimately” (Tillich, 1952). The philosopher/theologian Søren Kierkegaard (1813–1855), often referred to as the Danish Socrates, places himself firmly in this maieutic tradition, and anchors this birthing of the moral self in a teleological context. Just as Socrates’s pedagogy of “paideia”¹⁰ challenged the status quo of ancient Athens, Kierkegaard’s revival of the maieutic practice aims to offer an “existential corrective” to the idolization of objective knowledge (Furchert, 2005, 2022). At its center stands a complex psychological process of objective and subjective reflection: the double movement of existential appropriation. The first part of this double reflection concerns our objective thinking about a particular truth, telos or value,¹¹ or the factual “what” of its content; the second reflection concerns the subjective “how” of that very content or how one comes to relate to it (Furchert, 2022; Kierkegaard, 1846/1992).

The concept of existential appropriation thus provides for a path from mere reflection about values to actually relating to it as a “truth for me,” by both adopting that truth or value and putting it into practice. We show in the Introduction (Section I.2) how Gadamer has described this application process as crucial for any context where professional (or ethical) knowledge needs to be applied to “the concrete case” (Gadamer, 1996). Bridging this gap between knowledge and application, what and how, is an act of constant evaluation, reasoning, and decision-making, translating what we know into doing, and what we do back into knowing, always for a specific and unique situational context (Gadamer, 1996; Huff & Furchert, 2023).¹²

Thus, the famous Kierkegaardian “leap” is not a leap “of faith” where one abandons reason but the leap with which the actor bridges the gap between knowing and

⁹Not to be confused with cultural appropriation, the superficial taking of ideas from a culture not one’s own.

¹⁰While claiming to honor the goal of the Greek ideal of elite human development through cultural experience (paideia), Socrates’s approach to questioning knowledge and virtue also radically undermined its assumptions, focusing on the crucial role of choice in the individual (Herdt, 2019).

¹¹In our discussion, we will be using “truth,” “value,” and “telos” interchangeably. As will soon become apparent, moral truths, the sort of truths we are talking about, are ones that can be objectively studied but also involve a call to be adopted, taken seriously, considered with regard to what it means “for me.” This is often what is meant by a value – it contains a necessary, normative, element of the “should.” And it shares meaning with the ancient Greek use of “telos” as an end goal that gives meaning to and shapes the process. This admixture of normative and objective is actually a universal characteristic of the good across all societies (Shweder, 2012; Shweder & Menon, 2014; Shweder & Power, 2013).

¹²The German philosopher and psychiatrist Karl Jaspers describes appropriation as the heart of what he calls *Lebenspraxis*, the existential practice of an examined life (Jaspers, 1938).

acting, the very definition of reasonableness in practice: a continual movement from one domain (objective reasoning) to another (subjective engaging) which stands at the heart of moral appropriation, or of what it means to put a value into practice. Appropriation in the existential sense does not claim a single leap, but a constant bridging of the tension between objective and subjective reflection, theory and practice, knowing a value and engaging it.

Existential value or goal appropriation is qualitatively different from knowing about the value, or from understanding that one's culture "has" a particular value, or from borrowing someone else's values, or even from studying a particular value and approving of it. Instead, it is a continuous existential act of being concerned for, focusing on, and being compelled by the greater good.¹³ And without this continuous commitment, there is no fuel to run the engine that might overcome the judgment-action gap we describe in Chapter 5.

The challenge here is to hold together both learning processes at the same time, the one that reflects on a truth or value per se and the one that reflects on the learner's relation to that truth. Both movements complement each other and make for what we call existential appropriation. Where there is no proper objective reflection, subjectivity becomes blind and easily captive to ideology or fecklessness. Where there is no proper subjective reflection, objectivity becomes lifeless and disconnected to daily living or dangerously biased in the conception of the good. The persistent hermeneutic cycle¹⁴ of objective and subjective reflection describes the tension the moral actor must hold together.¹⁵ It is how one moves from merely objective observation of ethical values to personal choice and commitment and thus to the place from which motivation to ethical action springs.

But there are many values that we are balancing against each other at any time. How do we choose among the values we might apply in any situation? Existential appropriation highlights the teleological task at hand, namely choosing the kind of telos or higher good that guides our moral formation, by ultimately putting our long list of relative values into place. Social science research on values has recognized this hierarchical nature of values (Ciecuch, Schwartz, & Davidov, 2015). The theologian Paul Tillich (1952), adopting Kierkegaard, speaks of "ultimate concern" to describe this central organizing commitment in one's life toward "what concerns us ultimately." Here the existential task goes beyond mere value accumulation by transforming or transcending our relative values for the sake of the ultimate, the "value of values."¹⁶ Thus, existential appropriation is itself moral action: it requires interior

¹³One can see here how Kierkegaard anticipated current psychological work on moral identity.

¹⁴Kierkegaard's description of this "double movement" is one progenitor of the hermeneutic philosophical approach (e.g. Gadamer's hermeneutic circle), which, like Kierkegaard, also draws from the well of Socrates.

¹⁵Kierkegaard's terms for this are "contemporaneity" or "simultaneity."

¹⁶In the practical context of professional design work, JafariNaimi, Nathan, and Hargraves (2015) refer to this as the "value of values" that guides design and propose "service to the situations of human life and living" as the overarching goal for design. We have written on this in more detail in the context of value appropriation in software design (Huff & Furchert, 2014, 2023).

processes of choosing and prioritizing, of acquiring and giving up, of emotional and motivational change according to what concerns us ultimately, and thus the re-sorting of all other values in relation to it. (See Chapter 5, Section 5.3.2 for a discussion about how complicated this sorting can be.) This re-sorting requires evaluating my attachment to certain values and my need to detach from others. I might realize that I am ultimately attached to something that is not ultimate, or that my ethical commitments are rarely leading to engaging them when I feel I should. In the Introduction we present Elizabeth Cheney's decision to serve on the Select Committee to investigate the attack on the US Capitol as an example of this re-evaluation process.

Thus, existential moral appropriation requires a constantly renewed change of heart. It recognizes the pull of the ultimate (or what one regards as ultimate), which cannot leave one unchanged. While a value merely informs but does not transform the individual, it does not leap into practice either. Hence, existential moral appropriation describes highly transformative interior actions: one strives toward the good (or evil) and at the same time is ultimately challenged and changed by its value.¹⁷ One not only learns something *by heart* but also takes it *to heart*. Metaphorically speaking, moral formation requires the "education of the heart," as Kierkegaard revived it for a postmodern audience, which can be understood as the intimate process of embodied learning that shapes our action in the world.

9.1.3 Monastic/Spiritual Formation

The formation of the heart metaphorically understood as the moral and spiritual center of a person is also a crucial theme in monastic and spiritual formation across traditions. Most of the major religions have traditions of monastic practice in which individuals undergo rigorous training in search of spiritual formation as it is viewed within that religion.¹⁸ Self-transformation is a central aspect of this formation, and moral formation is a foundational aspect of that. For example, Benedictine and Zen Buddhist monks have exchanged extensive inter-monastic visitations and they find striking parallels (and some interesting differences) between Benedictine spiritual practice and Zen Buddhist practice, particularly with regard to the lifelong formation of what the Benedictines call humility and the Buddhists call non-egocentricity (Augustine, 1986).

In these traditions, it is the setting aside of self-serving commitments that helps to produce an unalloyed approach to loving both the self and others, often embedded in a complex virtue ethics.¹⁹ For instance, one can see the importance of humility, the "queen" of all virtues, in an early morality play by the Benedictine Abbess Hildegard

¹⁷For Kierkegaard, thoughtful objective reflection and subjective appropriation (what he calls "double reflection") should prevent inappropriate surrender to autocracy or ideology. Objective reflection necessarily involves stepping outside of motivated reasoning. The process is supposed to allow the reflector to distinguish relative values from ultimate ones. In practice, it likely involves some of the skills and attitudes we include under evaluative/self-transformative moral formation in Section 9.2.

¹⁸Excluding Rabbinic Judaism, which views living in family and the world as itself holy.

¹⁹We need to recognize that the word "self" is being used in multiple ways here, and that the various religious traditions themselves disagree on what might be meant by "the self."

of Bingen, *Ordo Virtutum*. Here, the virtues are personified as ideal archetypes or competencies toward which the self strives (von Bingen, 1141/1982, 1141/1990).²⁰ Constant practice in humility is a central aspect of most Christian approaches to monastic formation, as is non-egocentricity in Buddhism. The virtues a disciple in Advaita Vedanta Hinduism needs include non-attachment to the self and world (Rambachan, 2006); and it is only by achievement of this non-attachment that one can truly see oneself and others and have compassion. In western Africa one of the most active communities of Muslim Sufism, the Divine Flood Sufism of Ibrahim Niasse, similarly emphasizes the renunciation of the old self (Z. V. Wright, 2015) in order to love others. All these approaches see the development of a non-egocentric self as an ongoing requirement for the development of a broad set of virtues that themselves are seen as merely foundational for further spiritual progress.

These parallels across major religions suggest that there is something in the monastic journey that requires at least an instrumental goal of establishing a non-egocentric self, a self that strives first toward a self-transcendent goal, in order to achieve a healthy relationship with the divine and the world. In all these monastic traditions, humility, or non-egocentricity, allows one to truly see/evaluate the self, and to truly have compassion for others. In most discussions of monasticism, these goals are explicitly moral in nature.

One might think of humility as simply an emotion, as do Weidman, Cheng, and Tracy (2018) who use methods to map the structure of words associated with humility. Monastic traditions think of it more as a combination of skills one practices, attitudes toward the self and others, and appropriate self-transcendence. Weinstein, Przybylski, and Ryan (2013) propose three integrative characteristics of self-regulation that sound suspiciously like the understanding of humility present in our traditional accounts: (1) conscious access to one's emotions, motivations, and values; (2) taking responsibility for one's emotions, decisions, and thoughts; and (3) non-defensiveness in response to challenge.

The deep monastic traditions across various major religions have also supported a blossoming of spiritual formation and spiritual care approaches intended for those outside the monastic life. The voluminous writings of Thomas Merton (1948/1998, 1964, 1966, 1975, 2009) are an example of one attempt to make monastic approaches available more widely. The prolific work of the Buddhist monk Thích Nhất Hạnh has been widely adopted in the mindfulness movement in psychology (Shapiro & Weisbaum, 2020). Spiritual formation/direction for laypeople in the United States and Europe has been predominantly Christian in its emphasis, and uses techniques and principles from Christian monasticism (e.g. the Jesuit Spiritual Exercises (Dyckman, Garvin, & Liebert, 2001; Houdek, 1996)). It has more recently adapted approaches from the psychological renaissance of mindfulness (Allen & Fry, 2022). And the approach is now being expanded to interreligious approaches seeking common ground across major religions and secular spirituality (Addison, 2000; Cline, 2018; Plante, 2021; Ruffing, 2002).

²⁰For an introduction to von Bingen's philosophy, see Furchert (2019). See also Michael Casey's (2005) *An Unexciting Life* for a review of the complexities of the several models of monastic formation. Humility plays a central role here, and in monastic formation in most religions.

Both monastic and spiritual formation traditions contain aspects of the previous two approaches: *Bildung* (Section 9.1.1) and existential appropriation (Section 9.1.2). Each emphasizes some form of humility or non-egocentricity in concert with other values and virtues, and an ongoing process of inquiry and practice as central in the process of moral formation.

9.1.4 Themes and Questions from the Traditions

In the following we extract three aspects or themes that all the traditional accounts seem to share, which in turn help us propose a framework for our further investigation of moral formation:

1. *Constant evaluation and practice:* All three traditions see self-transformation as more than change in a cognitive or emotional structure (though they all expect at least this). They also emphasize the requirement for continual practice, for action that embodies the virtues or moral goals they seek. *Bildung* is an intentional, guided process of encounter, self-evaluation, and transformation. Existential appropriation is a continuous process of putting one's ultimate commitment into both interior and exterior action. And the monastic tradition involves regular consultation with a teacher or elder in a constant practice of moral and spiritual virtues. All three traditions emphasize careful evaluation of the goals and actions and the role of openness to change based on evaluation of one's practice. A corollary is that much like some models of expertise, though there may be some dramatic moment of turning, there is no single "leap" or conversion moment that produces the change, nor is there a final stage where one has arrived. Rather, one always continues in practice and development.
2. *Transformation of heart/personality/moral identity in service of a self-transcendent goal:* All three traditions see the appropriation or adoption of some kind of self-transcendent goal as central in the process. *Bildung* seeks transformation by encounter with the ideal in God, other people and cultures, the humanities, or art and nature. Existential appropriation grapples with the teleological task of choosing one's ultimate concern and putting this into practice by transcending other, lesser values or projects. Monastic and spiritual formation across (and without) religions requires transformation of the self in encounter with God or the spiritual. Non-egocentricity plays a central role in this process because the process consists of turning away from a self-serving goal and toward a self-transcendent goal. All the traditions see this transformation as anchored both in action/practice and in change of heart, personality, or moral identity.
3. *Schooling as an institutional or educative process that supports moral formation:* All three traditions view moral formation as some kind of encounter with something outside the self that spurs, guides, and structures moral formation. One among the many forms of *Bildung* was an explicit renovation in higher education in Germany. Monasticism across all religions involves the construction of moral ecologies that support the study and practice of moral virtues. Existential appropriation sees the maieutic, midwifing pedagogy in personal education as a central support for the individual, though it is suspicious of institutions.

We use these three themes to develop a framework for three aspects of moral formation: *evaluative/self-transformative* (Section 9.2), *normative/self-transcending* (Section 9.3), and *facilitative/educative* (Section 9.4).

- The evaluative aspect focuses on intentional encounter, self-transformative practice, and evaluation. We ask: *How does the individual form and evaluate their progress (or failure) toward the moral goals and commitments they have set?*
- The normative aspect focuses on the moral goals/values themselves and on the complex processes of appropriating them, of incorporating them into the self, and thereby transcending the self. These are all processes that have often been overlooked in current research. We ask: *Within what value space is moral formation occurring? And how does the individual actor change their relation to these values in a way that might be called moral formation?*
- Finally, with facilitative/educative aspects of moral formation we return to moral ecology and ask: *How can moral ecology support (or hinder) evaluative and normative moral formation processes?*

Each aspect of moral formation has its own complexities, which interact with each other in many ways and with the surrounding moral ecology. Distinguishing between them allows us to inquire how the aspects overlap, and to integrate many seemingly unrelated literatures. The proposed themes will also help us integrate what has been shown in other chapters into a more focused picture.

9.2 Evaluative–Self-Transformative Aspects of Moral Formation

Evaluative moral formation marks those skilled processes by which we evaluate and attempt to modify our moral actions in light of our moral goals or values. It thus involves recognizing relevant situations and internal states and processes, evaluating them, and making and implementing plans for change in a desired direction. We sketch here an overview of this aspect of moral formation as it has been covered in Chapters 4, 5, and 6. We then consider the narrative level of evaluation, which is not well covered in those chapters.

- In Chapter 4, we cover the recent turn toward idiographic models of personality that emphasize how the individual weaves their own path of moral consistency (Cervone, 2021; Mischel, 2004; Wrzus & Roberts, 2017). These models highlight the individual ways people *adapt* to their perceived environments and *integrate* a variety of influences in thought, feeling, and behavior, all in a way that allows for a *pluralistic* understanding of moral personality: that there is no singular moral personality (Wrzus & Roberts, 2017).
- In Chapter 5, we suggest three integrative characteristics of self-regulation that might support appropriate reflection on moral failure:²¹ (1) conscious awareness of

²¹And that look suspiciously like the virtue of humility. See also Eskreis-Winkler and Fishbach (2022) for a review of the emotional and cognitive barriers to learning from failure.

one's emotions, motivations, and values; (2) taking responsibility for one's emotions, decisions, and thoughts; and (3) non-defensiveness in response to challenge (Weinstein et al., 2013). There, we also accumulate the evidence of the conceptual complexity and flexibility of self-evaluation that make awareness, responsibility, and non-defensiveness important virtues in self-reflection.

- In Chapter 6, we review a useful model of personality change that is particularly relevant for evaluating moral formation (Wrzus & Roberts, 2017). The model concentrates on the awareness and evaluation of one's reactions to a situation and using this evaluation to construct *desired* expectations for behavior in similar situations in the future. These desired expectations can be evaluated when those situations arise again. There we also cover other skills relevant to evaluative moral formation, like self-regulation, moral attentiveness, and moral imagination.

9.2.1 Telling Stories

But there is another level at which evaluative moral formation occurs: the narrative. These skills, processes, and influences just mentioned deal well with moral formation focused on high-level traits (e.g. neuroticism) and mid-level projects and personality difference (e.g. cynicism) (McAdams, 2013). They do not say much about how we tell stories to ourselves about ourselves and how those stories might influence evaluative moral formation. These stories are often designed to connect our past, present, and future in a coherent manner. Stories allow us to see meaning in our lives and guide our lives as we navigate into the future (Bluck et al., 2005; McAdams & Pals, 2006). This overarching narrative, woven from a sequence of stories, hopes, and regrets about the self, can be one of the formative influences in a reflective process that guides and monitors moral decisions and plans.

Autobiographical memory, and the narratives we construct from it, have a significant directive function in our lives. These narratives help maintain personal identity over time (exerting some consistency pressure) and guide the planning and nurturing of current and future social interactions (another kind of directive influence) (Bluck et al., 2005; Rasmussen & Habermas, 2011). Narratives have been shown, for example, to be central in constructing a sense of agency for those in psychotherapy, with the agency in turn becoming a leading indicator of therapeutic benefit (Adler, 2013; Adler, Wagner, & McAdams, 2007). Thus, looking forward, we construct our agency in narrative and those narratives directly guide choice and planning (Bluck et al., 2005). Looking backward, we are especially motivated to make narrative sense out of negative occurrences and this reconstruction impacts our agency in helping others (McAdams et al., 2001).

How can narrative play a role as an arbiter in evaluative moral reasoning? "Telling a story" can have evaluative force because it contains autobiographical reasoning, the set of interpretive actions used to make sense of who we are, where we have been and where we are going. Indicators of structural quality in a narrative (indicative of coherence) include believable causal chains, reactions, and descriptions, finding growth or regression over time, connecting episodes with clear beginnings and endings, foreshadowing, and retrospective review (Adler et al., 2015). Other narrative content that can bring evaluative force on behavior includes themes of motivation (agency, autonomy, generativity, intimacy, etc.), affect (positive or negative tone, redemption,

contamination, anger, fear, etc.), and integrative meaning (change in moral identity, integration of contradictory themes) (Adler et al., 2015). The specific content of these themes in the narrative one constructs has implications for how the past and present are to be morally evaluated and how the future ought to be planned.

Is there any evidence that narrative content has implications for evaluative moral formation? There is at least clear evidence that this narrative aspect helps to distinguish moral exemplars – individuals who are widely agreed to be admirable moral models – from those whose lives may be conventionally moral but are not centered on morality. And there is much work that suggests that narratives are central in guiding the moral formation process in moral exemplars.

As part of an extensive research program that has compared moral exemplars to matched controls, Walker and Frimer (2015) did not find differences in personality traits or values, but did find differences at the narrative level. Moral exemplars' stories more often emphasized early secure attachment to parental figures, a lack of early enemies, and the presence of early support. In comparison to controls, their narratives (but not trait scores) had more positive affective tone, were more likely to positively appraise interpersonal communication, and had language that tended to emphasize the needs of others. Finally, the stories exemplars told also tended to see good coming out of bad occurrences (a redemption theme, McAdams et al., 2001). These individual stories are woven together into a narrative of the self with implications for current and future moral action (Lapsley, 2016): actions such as nurturing current social relationships and guiding the development of future relationships (Bluck et al., 2005). Colby and Damon (1992) see this embedding of the moral exemplar in relationships that guide their careers as a central aspect of the development of moral exemplars.

Still, there is very little work on the specific processes by which narratives guide evaluations of the self or constrain future moral action. The processes outlined by Wrzus and Roberts (2017) in Chapter 6 can surely be adapted to incorporate the narrative level.²² One can imagine additional ways narrative might be recruited in moral formation (e.g. see Section 9.4.3).

9.2.2 Critique from the Three Traditions

In sum, we have a vast and complex research literature, spread across several chapters of this text, on how individuals might evaluate their performance in achieving their moral goals. Some attempts have been made to integrate and unify these ideas (e.g. Wrzus & Roberts, 2017). But even these integrations do not give us the practical programmatic approach to moral evaluation that some of the traditional accounts do. Monastic and spiritual formation traditions have detailed and rich approaches to self-examination (e.g. the *dokusan* interview in Zen Buddhism or Jesuit *spiritual exercises* in Christianity). *Bildung* has produced detailed narrative accounts of its individual practice and programmatic revisions to higher education in the service of its goal. Most psychological research in moral psychology has yet to be connected to such practical applications in any systematic way.

²²As can the processes found in the traditional accounts, and those in clinical psychology, religious conversion, and volunteerism. The traditional accounts seem to fit much better to this narrative aspect of evaluative moral formation.

In addition, one can see from the tradition of existential appropriation that the distinction between evaluative and normative is not as clear as our presentation would have it. In fact, the evaluative necessarily leads us to the normative. It cannot be disconnected. Only a robust set of defense mechanisms, rigidly deployed, can keep us from evaluating whether we are in fact caring about the right things. The “what” of our goals and the “how” of our relation to them need to be kept in constant tension. This is part of the dynamic that Colby and Damon (1992) explore that draws exemplars ever further into broader commitment as they evaluate their current commitments. Psychological research on evaluative moral formation could benefit in many ways from conversations with these traditions.

9.3 Normative/Self-Transcending Aspects of Moral Formation

Much of what we call the psychology of moral action consists of processes (e.g. self-regulation, etc.) that can be analyzed in a content-free manner. That is, one can easily imagine a wide range of goals, values, or narratives that could drive these processes. For example, careful evaluative moral formation can serve a person’s goal of becoming more compassionate or ruthless in pursuit of a supreme value (Côté et al., 2011). Evaluative moral formation can be based on a narrative of good coming from the bad or a narrative of the constant, enervating resurgence of failure (McAdams et al., 2001). Moral attentiveness can be focused on threat from the outsider and compassion only for the insider (Hart, 2005b; Hart & Carlo, 2005). How can we find some structure in this confusing variation in values, ideals, goals, and projects that can direct both moral and immoral action?

9.3.1 Value Dimensions of Normative Formation

We have often mentioned pluralism in values in this book. What is the landscape of the variety of values from which people select some subset to guide their moral formation? If we want to become good, how do we determine what the good looks like? There are a wide array of works that identify dimensions of value in motivation content and little convergence among them.²³ There is even disagreement about the ultimate goal within and among the three humanistic traditions with which we began the chapter. Despite this disagreement and confusion, we will nevertheless attempt to systematize the variety of content and meaning that is relevant to moral becoming.

It is easy to multiply examples of the different content and values that can drive moral formation. It is less easy to bring some structure to the diversity in a way that helps us understand it. In this section we will, in the Humboldtian spirit, be looking for the diversity and the structure in this wide variation. What are the various kinds of values, goals, projects, etc. that motivate moral action?²⁴ Can we organize them in some way that allows us to understand their interrelationship?

²³This is in part because there are so many options and so little theoretical integration. See n. 5 for an incomplete listing of approaches and systems.

²⁴To see a more in-depth discussion of values as self-transcendent or ultimate, see the Introduction and Chapter 5.

All developmental moral psychology presumes some movement in values (e.g. from self-centeredness to the inclusion of others). It is implicit in the notion of development or formation that it is not mere change but movement toward some ideal or telos (Schachter & Ben Hur, 2019; Schnitker, King, & Houlberg, 2019). If we construct a space of values, goals, or projects relevant to moral action, how might the choosing individual move among them in commitment in a way that could be called formation?

Thus, the questions we ask in this section are: *Within what value space is moral formation occurring? And how does the individual actor attach or change their attachment to these values in a way that might be called moral formation?*²⁵ To engage these questions, we will look at five dimensions of the landscape of goals and values toward which one might orient moral formation, and from them construct a complex-enough three-dimensional space within which it would be productive to track moral formation and change.

9.3.1.1 Self-Transcendence In the Introduction, we proposed an existential turn (Section I.6.1) that reminds us of the need for self-transcendence to move from simply knowing that a value, ideal, or project has value to actually appropriating that value in one's own life. It is how one moves from merely objective observation of ethical values to personal choice and commitment, and thus to the place from which motivation to moral action springs. We argued there that this turn opens up the self to transcendence to the variety of goals, values, or projects that might be treated as ultimate by the individual. This is a move away from attachment to self-serving values to ultimate attachment to the greater good, or to ultimate values.

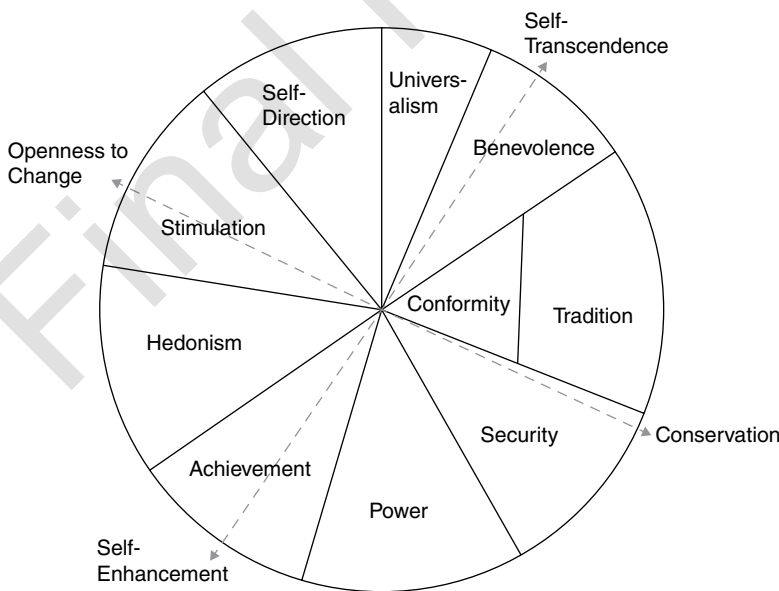


Figure 9.1 Schwartz circle of value regions.

Source: Adapted from Borg, Bardi, and Schwartz (2015).

²⁵These two questions mirror Kierkegaard's distinction between the *what* and the *how* of existential appropriation.

In Chapter 3 (Figure 3.1) we introduced a cross-culturally valid model of values that sees them as arranged around a circle in opposing pairs, with values near to each other on the circle being more compatible (reproduced here as Figure 9.1 for reader convenience). In it, one can see that *self-enhancement* (lower left) is opposed, across the circle, to *self-transcendence* (upper right). And at a near right angle to this is a dimension ranging from *conservation* (right) to *openness to change* (left). This is a second dimension in the value space that we will discuss in Section 9.3.1.2. The two dimensions form a circle with more specific values arranged around the circumference. The approach is based on the idea that cultures and individuals find themselves somewhere within the circle at a balance point between the various opposing values, endorsing all to some extent, but favoring some over others.²⁶ The circle provides a structure for the characteristic ways that cultures and individuals balance these values (Borg, Bardi, & Schwartz, 2017).

In Section 9.3.3, we will review extensive work by Walker and Frimer suggesting that there is interesting individual development over the life span on the dimension of self-transcendence, e.g. not just a simple movement from one to the other but an integration of the two ends of the dimension, with self-enhancement serving the goal of self-transcendence (Walker & Frimer, 2015). This crucial developmental sequence makes it clear that self-transcendence will be important to include in any model of moral formation. We cover this dimension in greater detail in that section. We label it as self-transcendence in part to connect it to other literatures (Frankl, 1997; Koltko-Rivera, 2006; Maslow, 1969) and in part to allow Humboldtian space for the wide variety of values or projects for which people self-transcend.²⁷

The self-transcendent end of this dimension needs to be considered at the idiographic level, because it is open with regards to the particular goal or telos of self-transcendence that each person might choose. See Section 9.6.1.5 on specific commitments and projects to get some idea of the diversity we are including here. Here, self-transcendence is more the task or process required than the end goal itself, which is complex and widely variable. Further, for each individual self-transcendence may be more complicated than one goal. Here, the complexity of the self and its commitments that we outline in Chapter 5 becomes evident. One might have moderate

²⁶This is a pictorial way of interpreting the rank ordering of the various values by different cultures (Schwartz, 2016; Schwartz et al., 2012). Empirical reality is a bit more complicated. Rank orderings in some cultures do not preserve the order of the circumplex (Schwartz, 2006), and thus the geometric balance point interpretation we offer here founders in these instances. The recruitment of one value in the service of another that we explain in Section 9.3.3 may help to understand some of these inversions.

²⁷Names of dimensions like these are always negotiable. Walker and Frimer (2015) replace “self-enhancement” with “agency” and “self-transcendence” with “communion” as their labels for the endpoints of this dimension, in part to connect it to other literatures. We prefer self-transcendence in order to emphasize the multitude of non-social goods to which one might be devoted (religion, art, environment, etc.). See the discussions in the Introduction and in Chapter 5. It is important to note that the label “self-transcendence” here is not in itself a value, goal, or project. It is, rather, a stand in for the multitude of goods for which individuals might self-transcend and that they might take as ultimate. This fundamentally alters the dimension by not anchoring this end in pro-social values, while still allowing them among other ends.

self-transcendence toward the value of compassion (producing a general pro-social attitude) and passionate, even ultimate, dedication to family or some other life project. These values might be integrated in some way (e.g. with compassion as the theme in a particular narrow project) or exist in tension across separate domains of the multidimensional self (e.g. work projects vs. family commitments). And to the extent that a particular value or goal is really considered to be ultimate, there should be tension to integrate it across the separate domains of the self. The point of ultimate value is that for some people, at least in some domains of their life, some value, telos, or life project is taken as the most important thing. And for some people, this extends over all domains in their life.

We must include a caution here on recognizing the multiplicity of possible goals in self-transcendence. Some goals for which people sacrifice themselves involve causing harm to others. Koltko-Rivera (2006) grapples with this difficulty in his review of Maslow's understanding of self-transcendence and admits that terrorism and other forms of evil can, at the idiographic level of personal commitments, look a great deal like self-transcendence. Thus, commitment to a self-transcendent goal has a dark side (Hart & Carlo, 2005; Skitka, 2014; Skitka & Mullen, 2002; Washburn & Skitka, 2017). Scholars will need to do the work to distinguish between the goals. It may be that the psychological processes are very similar if not identical. Perhaps some of the other dimensions we list in Sections 9.3.1.2–9.3.1.5 will help us do so (e.g. inclusion). Opening up this dimension to this puzzle of identifying appropriate ultimate goals allows us to frame the problem of the crucial and unavoidable work of distinguishing what the good is and whether how we become good is the same process as how we become evil. This brings us back to the idiographic struggles of each individual as they engage in their teleological task of contemplating their ultimate concern (see Section 9.1.2).

9.3.1.2 Openness to Change The second dimension in the figure is *openness vs. conservation*. This seems to be treated more as a personality characteristic that does not change much over time. Research seems to focus mostly on the ways this dimension tracks the (supposedly static) liberal – conservative continuum in the United States. But we might also find some interaction between trajectories in the two dimensions. Those with conservation commitments might show a different pattern of development from self-enhancement to self-transcendence than those with commitments to openness. This is speculation at this point, but one can imagine those who mature on the self-transcendence dimension in politically or institutionally or culturally conservative households might select very different kinds of life projects than those who might mature in households that value political, institutional, or cultural change. Seeing these two dimensions as forming a moral formation space allows us to ask these questions, and again to expand the space of what is moral and how people choose to be moral.

9.3.1.3 Inclusion/Exclusion This dimension is the extent to which we include others in our understanding of those deserving of moral consideration. The move from self-enhancement to self-transcendence does not remove us from the messy difficulties of deciding *for whom* one should have compassion, or how others are included in one's religious commitments, or for whom social justice is important. How wide should

one's circle of concern be? We should note that the value circumplex model has "universalism" as one value at the self-transcendence end of the dimension. And this suggests that the more one endorses or appropriates communion the wider one's circle of concern should be. But the Schwartz (2016) model is based on averages across many people, and there may well be those for whom self-transcendence does not entail opening wide the gates to all.

For instance, we already know there are sub-types of moral exemplars. The distinctiveness of two of the sub-types (helper vs. reformer, see Chapter 3, Section 3.2.1.4) consists in part in *for whom* they are moral. Helpers care for certain individuals with needs, and do not much concentrate on societal change. Reformers want to change systems as an instrumental goal in service of producing justice or care for others.²⁸ The latter may be more universalist in their embrace of others, while some helpers may be more restrictive in their compassion. The "deserving poor" is a standard trope for those who want to limit social assistance. Still, even reformers limit their focus to particular social justice issues, and thus leave out those affected by the neglected issues. A more theoretically targeted selection of moral exemplars might help us grasp some of the variation in ways that people do good and the extent of inclusion of others in that moral action.

Susan Opatow explores this dynamic of with whom one values communion under the label moral inclusion/exclusion (Clayton & Opatow, 2003; Opatow, 2001, 2005) and we adopt this language for this dimension. She has reviewed how moral exclusion is constructed in a moral ecology (Opatow, 2001) and how programs stressing moral inclusion can help to reduce violence and injustice (Opatow, 2005). Others have begun to investigate what they call moral expansiveness, the inclusion of a wide range of others in their circle of concern. This work has shown expansiveness to uniquely predict willingness to prioritize humanitarian and environmental concern over personal and national interest (Crimston et al., 2016). Similarly, Graham et al. (2017) propose centrifugal and centripetal forces and frame them as separate and competing pressures to form some balance along a continuum they call the moral circle. And, of course, work on altruism that used in-depth interviews with those who helped Jews escape the Holocaust found a personality characteristic called extensivity (the extent to which one included others in the circle of moral concern) to best differentiate between those who helped and those who did not (Einolf, 2010; Oliner & Oliner, 1988).

These parallel theoretical accounts suggest there is likely another important dimension along which normative moral formation can take place. One might have great integrity about justice and care for those we include in the deserving community, while being callous and uncaring for those "other people." We show a different moral face to those who are "them" than we do to those who are "us." One can also limit one's moral horizon without overt prejudice, by unexamined or intentional parochialism. Again, the Munsons (see the Introduction) serve as an example. At what point in their progression of commitment do they become exemplary? Even at the beginning,

²⁸One can, of course, do both or neither, as do some of the exemplars from Huff and Barnard (2009). Those doing neither were focused on more specific projects like mentoring (see Section 9.3.1.4 for what specific projects might look like).

when caring only for their own foster children, they are self-transcending, but their self-transcendence is inclusive of only a small circle.

We know a little about what might influence the inclusion dimension relevant to normative moral formation. It seems to be flexible depending on moral framing and domain but also reliably related to the development of moral identity.²⁹ Its flexibility is influenced in part by empathy. In everyday life, we primarily experience empathy for close others rather than those at a distance (Depow, Frances, & Inzlicht, 2021). However, empathy is not inherently exclusive and one can frame judgments in a way that fosters egalitarian empathy (Fowler, Law, & Gaesser, 2021). One can likewise frame judgments in ways that encourage “virtuous violence” against the other (Cohen-Chen, Pliskin, & Goldenberg, 2020; Fiske & Rai, 2014). Cultural preferences for tight framing of moral constraints and the moral circle can anchor inclusion but also change it as the culture reacts to external threats (Gelfand, 2012, 2021; J. C. Wright, 2021).

There is some evidence that higher levels of moral identity are associated with a wider “circle of moral regard” (Aquino et al., 2006; Reed & Aquino, 2003). And even for those who value group loyalty, commitment to moral identity reduces the extent to which they are willing to endorse torture in its service (Smith et al., 2014). This at least allows us to say that even though conservative and liberal value commitments might look different, if they have integrated morality into their identity one can expect they will share some values (Frimer, Tell, & Haidt, 2015). So, we may need to differentiate between inclusion with regard to basic humanity (e.g. appropriate targets for compassion and human rights) and an inclusion that involves a more complete reception, incorporation of, and obligation to the other. Beyond simple viewing of others as part of “us,” participation within a community helps to construct our sense of self. Seeing one’s self as socially connected influences compassion and socially responsible behavior (Cojuharenco, Cornelissen, & Karelaia, 2016; Cross, Bacon, & Morris, 2000; Day & Impett, 2018; Vignoles, 2018).

In a longitudinal study of high-school students, Pratt et al. (2003) found that self-ideal (measured by self-report of traits one strives for, such as being a good citizen, fair, just, caring) predicted community involvement, but over time it was community involvement that led to increases of these self-ideal traits. Thus, active participation in community can produce a virtuous cycle of self-ascription of traits, followed by further community service. Hart (2005a, p. 260) argues that “If the notion of identity is to contribute to an understanding of moral functioning, then it must be a construct with deep roots in a social world.”

As we consider incorporating inclusion/exclusion in our moral space, these roots will need to be both deep and specific to each individual and his or her personal history and moral ecology. Piliavin and colleagues (Grube & Piliavin, 2000; Piliavin & Callero, 1991) have shown the importance of “specific role identity” (as, e.g. a blood donor or a cancer society volunteer) in supporting volunteer activity. Rule and Bebeau (2005) and Huff and Barnard (2009) have also tracked the importance of specific

²⁹Graham et al. (2017) provide a short history of the use of this dimension and survey a host of factors that can influence one’s position on the dimension.

identity themes in professions. The extent to which these specific commitments limit or facilitate moral inclusion is as yet unexplored.

9.3.1.4 Rules vs. Goals We also need to consider the possibility that, for many people, or even for most people at one time or another, morality is not really about the pursuit of moral goals but instead the following of moral rules that serve as guardrails.

Most professional codes of ethics include two types of items: things one ought not to do, and moral ideals one should approach (Harris, 2008). In Chapter 6, we explored Kuhl and Koole's (2004) distinction between the self-control system and the self-maintenance system. This is roughly the distinction between those processes that help the person focus on the task at hand and avoid temptation and those processes that direct the person to take valued action.

In the domain of engineering, the philosopher Charles Harris (2008) presents a parallel distinction between *preventive* values and *aspirational* values. Preventive values are about "not violating moral rules." These constitute about 80% of the text of the National Society of Professional Engineers code of ethics and are essentially negative in form (e.g. not approving documents that do not conform to standards). Aspirational values are about approaching moral goals or values. Harris (2008) gives examples of designing safety into engineering systems that goes beyond codes and standards, evidencing a deep commitment to safe engineering, or to environmentally thoughtful engineering, or to engineering that takes its social embedding and impacts seriously.

Similarly, Janoff-Bulman and Carnes (2018) use the basic motivational processes of approach (prescriptive) and avoidance (proscriptive) to frame a model of moral motives that operates at three levels, the intrapersonal, the interpersonal, and the collective. This taxonomy provides useful insight, for instance, into the differences one finds at the collective level between conservatives, (who favor social order – an avoidance goal) and liberals (who favor social justice – an approach goal).

Thus, the domain of moral rules and obligations (e.g. do not lie, cheat, plagiarize) may well be associated with the self-control system and have its own trajectory in moral formation. And the domain of aspirational goals (e.g. be a blood donor, care for the poor) may be associated with the self-enhancement system, particularly when these moral goals are an important part of the self (Cervone & Tripathi, 2009).³⁰

Additionally, each system may have its own distinctive pull on the moral action of the individual. As Nucci (2004) notes, even if individuals do not have morality as a defining aspect of the self (and are thus not engaging the self-enhancement system in the service of moral goals), there is still something about the moral domain that makes it difficult simply to ignore rules. "I'm not good at academics," sounds qualitatively different (and much less dangerous) than "I'm not good at morality." We expect the pull of rule-based, inhibitory, preventative morality to be widely shared, even by those for whom moral goals are not central to the self.³¹ But there is very little work on how these two systems interact or how the two different kinds of moral motivations differ.

³⁰See Chapter 6 for more on this distinction.

³¹But we make exceptions for (and even inversions of) those rules when it comes to treating outsiders and enemies (Cohen-Chen et al., 2020; Fiske & Rai, 2014).

9.3.1.5 Specific Commitments and Projects So far, this discussion is operating at a highly abstract level of general themes in morality (self-enhancement and transcendence, moral identity, moral exclusion, rule following) rather than the concrete particulars of the specific causes to which people normally dedicate themselves.³² It seems likely, however, that the process is not one of becoming good at this abstract level, but more one of becoming good by being attached to a specific, concrete, personal moral goal (rather than an abstract value like communion, or even more vague, self-transcendence). Across all the research on exemplars, moral commitment appears to be particular. Despite Suzie Valadez speaking of God's call to love all people, the people she serves are the poor "people of the dump" in Ciudad-Juarez, Mexico (Colby & Damon, 1992, chapter 3). Even though Stephen Engberg is committed to privacy rights and speaks at this abstract level, the particular privacy he cares about is in the design of business software in Denmark (Huff & Barnard, 2009). Though Walker and Frimer (2007, p. 848) identified value dimensions behind particular concrete commitments, their exemplars were selected because they showed concrete "extraordinary and long-term commitment in providing care to individuals or groups, [or to] community service or humanitarian causes." How do we get from smaller, concrete commitments to larger, more abstract ones? Colby and Damon (1992) propose a process that might lead from local, particular moral commitments to broader commitments and offer several examples of that process in their exemplars. They call these *environmental personal growth affordances*. But only a few of their exemplars show a clear pattern from specific to general.³³ The more common development seems to be from lower to higher levels of commitment within a specific domain.

In a similar vein, Koltko-Rivera (2006, p. 302) explores Maslow's concept of self-transcendence as "a motivational step beyond self-actualization." It is also clear here that the motivation is particular, as self-transcendence may involve "service to others, devotion to an ideal (e.g., truth, art) or a cause (e.g., social justice, environmentalism, the pursuit of science, a religious faith), and/or a desire to be united with what is perceived as transcendent or divine" (Koltko-Rivera, 2006, p. 303).

A similar pattern can be seen in research on how participation in volunteering or social activist work affects those who volunteer. A classic example is the study by McAdam (1989) on the effects of the 1964 Mississippi Freedom Summer project on 212 participants in the project, as compared to 118 who signed up, were accepted, but did not go.³⁴ Compared to the no-shows, the participants were still deeply engaged in politics twenty years later, were more likely to be unmarried, have lower income, and to be concentrated in the helping professions (e.g. teaching). The study

³²Frimer et al.'s approach to categorizing agency and communion as terminal or instrumental involves either interviews in which people must identify values as terminal or an elaborate coding approach that finds the abstract themes in the text of the interview (Frimer et al., 2011). It is at this level of abstraction that they find patterns. In most of the interviews, their participants speak movingly about commitments to specific projects and actions.

³³We know of no other systematic data on this trajectory or its influences.

³⁴These were volunteers from all over the United States who came to Mississippi for a summer focused on civil-rights organizing, non-violent demonstration, and civil disobedience in support of civil rights for African Americans, particularly in the Jim Crow south.

is based on the completed applications of the two groups (obtained from archives) and a follow-up survey sent to them in the late 1980s. The no-shows “did not differ significantly from the participants in the values they brought to the project ... [and] appear as essentially alike on a list of variables: race, social class, type of neighborhood, home region, type of school, and major in school” (McAdam, 1989, p. 749) But participation in the Freedom Summer likely affected their self-concept and their later work (and even romantic partner) choices. The intertwining of personality, moral identity, skills and knowledge, and moral ecology are quite complicated here (see e.g. Bosi, 2016; Giugni, 2004; McAdam, 1989) but the effects of immersion in the complex moral ecology of civil resistance seem undeniable. By choosing to dedicate their summer to this project, the participants instigated a change in the national consciousness but also in their own.

Abstract level commitments can also spawn particular concrete commitments. Sophie Scholl was committed to the White Rose resistance movement because, in part, she was committed to exposing the injustice of the German regime in WWII (see Introduction, Section I.3 and Coda). The more general commitment to support oppressed groups that one finds on the left in Western countries has been shown to be a source of extreme commitments to particular outgroups (e.g. Palestinians, Kurds) based on the perception of the group as unjustly oppressed. In some cases, identification with the oppressed out-group (e.g. Kurds) is even stronger than identification with their own in-group (e.g. Canadians), producing an interesting inversion of the inclusion dimension (Kunst et al., 2018).

It seems likely, though, that most people with extraordinary levels of commitment to a particular cause did not decide to “become good” at a general level but instead had some initial attraction to a particular project (perhaps articulated as a more general ideal) and then became more committed. One may not plan to make high levels of commitment to a cause, but one could instead stumble upon or be drawn into it. This attraction itself might be based on previous value commitments that were free-floating and unattached to any specific cause. But there is very little research that would help us untangle the precedence of commitment to abstract values vs. specific causes.

The philosopher Gadamer’s work on the problem of application (Gadamer, 1975, 1996) suggests that specific commitments in applied domains are unlikely to be the simple “application” of general commitments and principles to that domain. Moral action at the level of commitment to these particular causes (Palestinians, Kurds, civil rights, etc.) or domains (e.g. engineering, medicine) is often thought of as simply applying knowledge (e.g. about rights) to specific cases (e.g. civil rights). But careful analysis of the problem of understanding in “application” suggests that there is knowledge and truth to be discovered in the process of application itself, knowledge that is unavailable from abstract knowledge of the domain. Knowing about privacy rights, for instance, is only a very small part of what one needs to “know” to be able to design software that protects privacy (Huff & Furchert, 2014). Movement back and forth from a specific application area to general principles seems more likely than top-down application. Goal-directed moral formation does not have to be an entirely conscious process. Our goals can guide the focus of our attention, and this influence can happen outside of conscious awareness (Dijksterhuis & Aarts, 2010). This influence is particularly evident when those goals are ones that have been practiced to the point that pursuit of them has become habitual (Aarts & Dijksterhuis, 2000).

As a part of the general pattern of development from agency to communion, there is clearly an increase in commitment to communion values from childhood through early adulthood. At this point, we know only that this trajectory has been documented. We have no evidence that it is the *only* trajectory that one would want to call moral formation, and some evidence that alternative paths exist. We know little about what individual differences or moral ecological influences there are on the trajectory. And, except among moral exemplars, we know little about the trajectory or its variation or influences among adults and older adults.

9.3.2 A Three-Dimensional Moral Formation Space: Self-Transcendence, Openness, and Inclusion

We have just reviewed five dimensions (self-transcendence, openness to change, inclusion, rules, and specific projects) on which one might look for variation in values and commitments. One implication of this review is that there is a need for thoughtful integration of the possible dimensions of the value landscape in moral action. With this tentative map of the landscape, we can document some movement within it that might count as normative moral formation – the choosing and appropriation of moral goals and the process of putting the normative into action. The appropriate criterion for success of any dimensional model is not whether the dimensions are “real” but whether they are helpful in discovering patterns and variation in moral formation and change. Our interest here is identifying a manageable subset that allows us to grasp the complexity of the differences in the good, or in moral commitments.

We begin by leaving aside two possible candidates. The space of specific commitments and projects (Section 9.3.1.4) is simply too diverse to provide much explanatory power at the nomothetic level of averaged tendencies among many people.³⁵ The rule–goal distinction (Section 9.3.1.4) is more of a dichotomy than a space within which formation might occur. And particular rules and goals are legion – see the lists of rules and goals in any professional code of ethics.

We are left with a three-dimensional landscape that we will lightheartedly label TOI: self-transcendence, openness, and inclusion.³⁶ The dimension of transcendence is about *what* goal is chosen, while the other two are focused on *how* and *for whom* that goal is approached or implemented. One can imagine (and there appear to be documented instances of) people moving along one dimension without much change on the other two.³⁷ There is a significant research literature associated with the moral

³⁵ One could reduce the complexity by forming types of projects (e.g. social justice, artistic, helping, etc.) in the same way one might think of types of goals for self-transcendence.

³⁶ Pronounced “toy.”

³⁷ Movement within this space might also occur at several different levels of personality. McAdams (2013, p. 276) borrows language from William James (1892/1963) to ask the question: “What does the I see when it looks on the Me?” He proposes three ways or aspects of how this might happen, actor, agent, and author. In terms of moral formation process we might call these *becoming good* (actor), *seeking particular goods* (agent), and *moral authorship* (author). This additional layer of complexity in moral formation provides even more richness in how people shape their formation and gives the opportunity for further labyrinths of self-deception or surprises of moral critique and insight.

aspects of each dimension. This alone argues that the dimensions in themselves are important for moral formation. And they do seem to be dimensional, allowing for various positions along a continuum (or perhaps multiple positions depending on the domain). Whether they are independent dimensions is a matter for empirical study, though work we reviewed in Sections 9.3.1.1–9.3.1.3 suggests patterns in which they may be somewhat correlated.

The point of constructing a dimensional model like this is to open up exploration in the variety of ways people make moral commitments and take moral action. It helps to structure and identify the different ways people might self-transcend. Those who transcend the self in service of a project might be committed (or resistant) to social change and shape their project or value commitments accordingly. Or they may be narrowly focused on helping a particular group, or motivated by the threat they perceive from another group. Exploring these variations in openness and inclusion will expand our grasp of the many ways people embody moral action.

9.3.3 Direction of Change in Moral Formation

In addition to variation across individuals in value commitment, one can find variation across time in a person's life. Schwartz (2006, Table 1) reports, for instance, that values of agentic self-enhancement decrease over the life span in a variety of cultures, while communal self-transcending values increase over the life span.

Over time, how might we become more involved in pro-social roles, or more committed to moral strivings and goals, or able to more thoroughly integrate themes of moral action into our life narratives? Can we even do so? Do most people's values "improve with age ... develop in the direction of greater tolerance, or wisdom or some other imagined enlightened state" (McAdams, 2015, p. 233)? There is surely a developmental progression in the moral domain from infancy to adolescence in terms of greater complexity and more pro-social orientation. Kohlberg's research program in moral judgment documented this, as has much developmental research since then (see Gibbs et al., 2007; Lapsley, 2020; Lapsley & Hardy, 2017 for reviews). But is there "improvement with age" over time for adults? Schwartz's (2006) cross cultural research program in values regularly shows increases in values of benevolence across the life span. However, beyond this documentation there is little systematic evidence. McAdams (2015, p. 234) answers his question thus: "Scientific evidence for such a romantic view is sparse."

In fact, looking for simple value change over time may run into the same complexities that at one time led social and personality psychologists to conclude that there were no stable effects of personality on behavior (Mischel, 1968, 2004). Recall the fruitless search for consistency in "honest" behavior in children across situations by Hartshorne and May (1928). Only after decades of research did more sophisticated efforts find complex patterns of consistency that differed from one individual to the next (Cervone, 2004; Roberts & Wood, 2006; Shoda, LeeTiernan, & Mischel, 2002).³⁸ A similar pattern may be emerging in normative moral formation.

³⁸See Chapter 4 for more on this.

In a groundbreaking series of studies, the Canadian psychologists Walker, Frimer, and their colleagues (Frimer et al., 2011; Frimer et al., 2012; Walker & Frimer, 2015) have been identifying complex patterns of value change across the lifespan that they argue can clearly be labeled as moral development. They have documented changes in value themes in adolescents to adults and linked these changes to either actual moral behavior (as seen in moral exemplars) or to self-reports of moral behavior. They are thus building a foundation for an argument about moral formation across the life span. They situate their research program within the value space provided by Schwartz's value circumplex. Their central finding shows a move from characteristic adolescent endorsement of *agentive* values (e.g. personal achievement and competence values, on the lower left side of the Schwartz circle) to *communal* values in adulthood (e.g. benevolence and universal care values, on the upper right side of the circle).³⁹ But this change is characterized by a certain complexity. The movement is not simply from endorsement of one value to the other but also involves the *way* the value is endorsed. As adolescents and adults make this developmental transition they move from (1) seeing agency as *more* important than communion, to (2) seeing them as *both* valuable, and then (3) transforming their valuation of agency by seeing it *as a means* to achieving communion. Thus, agency is at first a terminal value, valued for its own sake, but later becomes at least in part an instrumental value, valued because it is instrumental in achieving communion. Their evidence for this transformation comes from a series of studies done with adolescents and adults using both qualitative and quantitative approaches (Dunlop, Walker, & Matsuba, 2013; Frimer, 2012; Frimer & Walker, 2009; Frimer et al., 2011; Frimer et al., 2012; Walker & Frimer, 2015).

This research program suggests a developmental trajectory in the moral domain: that "displacing agency with communion as the terminal value should be a fundamental goal for moral development" (Walker & Frimer, 2015, p. 430). Note that with this claim, we have crossed the boundary from descriptive claims to normative claims: communion as the terminal value "*should* be a fundamental goal." This move from the empirical to the normative parallels Kohlberg's earlier claim that he had found an empirical foundation for a normative ethic (Kohlberg, 1971). This approach is a slightly stronger claim than Kohlberg's in that it is based in part on empirical observation of moral exemplars. This trajectory of displacing agency with communion may itself vary among individuals depending on what other value commitments they have.

For the sake of accuracy, we have presented the work of Walker and Frimer (2015) so far using the terms they prefer, *agency* and *communion*. But, given our Humboldtian approach, we have spent much of the book arguing that there are multiple goals, values, or projects one might call ultimate, and therefore anchor the moral realm for the individual who chooses that goal as ultimate. So, we have been using the terms self-enhancing and self-transcending to encompass all the possible values, goals, projects, etc. that people might take as ultimate concerns.⁴⁰

³⁹As we discuss this research program, we will use the original dimension names (agency vs. communion) from Walker and Frimer (2015), rather than our preferred labels of self-enhancement to self-transcendence.

⁴⁰In agreement with the dimensional terms originally used by Borg et al. (2015).

The heart of moral formation is this self-transcending value-goal-project appropriation process. It includes not just an internal reorganization of commitments but also skilled action, putting the self-transcendent commitment into practice. The psychological question is: How do these processes of commitment and enacting work? One answer is to look to social-emotional-reflective equilibrium (SERE).

One might think that this sort of embodied commitment to a self-transcendent goal conflicts with the idea of equilibrium that we have identified in Chapters 7 and 8 (among others). Self-transcendence, at least at its extremes, can produce extraordinary and life-altering commitment for some moral exemplars. But the two concepts are really just different aspects of the same process. Equilibrium in SERE does not mean finding the easiest path but the most appropriate and practically viable answer to the question: What does a person like me do in a situation like this? When we commit to a self-transcendent goal, we are still left with the everyday difficulties of implementing our goals and intentions in the historical situations in which we find ourselves. It is in fact the continued practice of SERE within ongoing moral action (and the continued practice of self-transcendence within SERE) that makes self-transcendent moral action possible.

9.3.4 Critique from the Three Traditions

The issue of inclusion, *for whom* moral identity is relevant, is rarely addressed in the moral psychology literature. Both *Bildung* and monastic traditions make it central. *Bildung* is at its roots aristocratic, but it can provide a platform to embrace wide cultural diversity. Monastic and spiritual formation traditions across religions also view compassion for the other as central to development, though they emphasize care for those “in the community.” These traditions can help maintain the complexity of the “for whom” when we want both to avoid “ethnocentrism” and care for those to whom we have special obligations.

And centrally for this normative aspect of moral formation, the hierarchical and ultimate nature of value commitments needs to be better addressed. Kierkegaard and his secular adopters are clear about the decisive role of the ultimate telos (as was his *Vorbild* Socrates with his stressing of the “highest good”) in the rearranging of one’s entire life (and not just one’s “values”).⁴¹ True encounter with and appropriation of the ultimate good is transformative to the individual. Surely not as a single experience but as an ongoing dialogue and reckoning as the phrase “ultimate concern” indicates. The Socratic method of midwifery is commonly thought of as based in leisurely Socratic discussion among aristocrats lying on couches, but the metaphors of giving birth and transformation of the heart make it clear that this is a much more extended process than an evening’s relaxed conversation, one that re-sorts numerous life priorities and painfully touches many aspects of one’s life (as images of being pregnant, in labor, and giving birth suggest).

One can see this re-sorting and recognition of value happening in many of our historical cases. Elizabeth Cheney, for instance, spent most of her political career championing traditional conservative values. It was not until she was confronted with

⁴¹ See, for example, Gadamer (1975) on the revealing and concealing of *truth* in art.

a transformed political landscape that she found she needed to insist on pre-existing deeper commitments to, for instance, truth-telling and the rule of law. This necessarily involved a painful reckoning with her previous commitments. She describes this ongoing adaptation based on ultimate commitments in an interview: “at each moment, ... I knew that I had to do what was right. And I knew that the obligation of elected officials to abide by our oath, has to be more important than party, than partisanship, or than any political office” (Karl, 2022 August 21). The intimate psychological details of this elaborate dance of attachment and detachment of relative and ultimate ends (which necessarily includes self-transcendence) are yet to be explored by many who possess the tools to begin to grasp it. But it will have to involve a focus on *how* one values, relates, commits, and attaches to ultimate and relative ends and how this dance is done in changing life circumstances.

9.4 Educative Aspects of Moral Formation

We have tried in various places to argue that the pull of the moral *should* is not always a social pull. It can sometimes be religious, or based in dedication to “truth, goodness, beauty, perfection, excellence, simplicity, elegance” (Maslow, 1969, p. 4). But all these *shoulds*, in order to be sustained, require a social matrix in which to thrive. All three of our formation traditions agree that there must, at the minimum, be something beyond the self that guides, encourages, and enables formation. Even the early Christian hermits living in the desert (the “desert elders”) required a social system to form their search for holiness (Stewart, 1991). We share our moral judgments with others and are influenced by the moral intuitions of others who are important to us. We look at moral exemplars, read books, attend religious services or theater, choose and gather friends, and go to schools to be inspired and to learn from them. These experiences can motivate us to change our intuitions, even without conscious reflection, and produce a virtuous (or vicious) cycle of influence in which our narratives or goals guide and nurture our social surround: and the surround in turn influences us at both implicit and explicit levels (Wrzus & Roberts, 2017).

9.4.1 Moral Formation as Planned Moral Ecology

We noted that a central theme in all three traditions is schooling, with two of the traditions dedicated to founding institutions to support moral formation. Wilhelm Humboldt’s idea of *Bildung* (Humboldt, 1792/1960) was institutionalized first at the University of Berlin (later named Humboldt University) and dedicated as least as much to realizing the potential of the individual as it did to learning concepts, ideas, and traditions. For this reason, it incorporated research (in the humanities and sciences) so that the individual could take a critical stance to tradition, at least in terms of seeing it as incomplete and open to revision (Herdt, 2019).

Monasticism across the religions is also an explicit attempt to create communities (even communities of hermits) to support moral and spiritual formation. The rules and structure of these communities vary, but they are all justified by, and connected to, their role in helping the community and the individual in that community to concentrate on their moral and spiritual tasks. This has resulted in spectacular

architectural achievements (e.g. the Taung Kalat monastery in Burma) and millennia-long continuous traditions of practice (e.g. those communities that follow the Rule of Benedict in Christianity), both in support of moral formation.

Kohlberg's educational program was headlined by the "Just Communities" approach, which also saw its role as creating communities where children, teachers, and administrators could learn to construct a more just and compassionate learning environment together, and thereby form themselves into more fair-minded and caring persons. The ultimate (and unfulfilled) goal was that the schools and their graduates would become transformative in society (Higgins-D'Alessandro, 2015). Kohlberg was clear that schools – all schools – teach a moral curriculum, whether it is hidden or intentional. His attempt was both to acknowledge this hidden curriculum and to be intentional about constructing it using what we know about moral formation.

The term "hidden curriculum" was first coined by Philip Jackson (Jackson, 1968) to describe those aspects of the curriculum that are implicit and at least in part unintentional in American K-12 classrooms. A significant part of the hidden curriculum is moral and focuses on what we should value and how one should act. It has become an accepted term of art in general education, as well as in medical education (Balboni et al., 2015; Lawrence et al., 2018). It does not take much imagination to extrapolate the hidden curriculum to the occurrence of socialization in organizations that makes corruption more (or less) likely. We discuss these processes in Chapter 3.

Last, but not least, the tradition of Socratic *paideia* (Section 9.1.2) has been influential in both educational and therapeutic settings, emphasizing the maieutic aspect of moral and individual formation (philosophical midwifery). Much of the existential tradition understands schooling of the heart to be best practiced outside institutions, either in dialogue with a trusted teacher or *Vorbild*, or in dialogue with literature, art, and one's own reflections. This process may be better represented in the psychological literature as social influence.

9.4.2 Moral Formation as Social Influence

Interaction with others who support or challenge our goals can shape both evaluative and normative aspects of moral formation. For instance, a central role of parents is to provide evaluation, support, and guidance to children in their development of moral character (Lapsley, 2020). Close friends, mentors, and spouses can support or inhibit our attainment of valued goals, particularly when they encourage goals that are congruent with our own ideals for ourselves (Rusbult, Finkel, & Kumashiro, 2009). People who are very good in self-regulation (e.g. in organizing themselves to achieve their goals) tend to shape their social surround to support their goal pursuits (Ent, Baumeister, & Tice, 2015). They prefer to spend time with, collaborate with, and be informed by others who are themselves good at goal pursuit or who are directly instrumental to their goals, choosing colleagues and mentors who support their goal pursuit (vanDellen et al., 2015). Though people with low self-control tend to reap the most benefit from social support for their goals, they unfortunately are less likely to find themselves in situations where they receive it (Nielsen & Bauer, 2018).

But beyond this, interaction can shape and change the moral goals we choose, the rules we emphasize, and the range of our moral concern. Colby and Damon (1992), in their classic study of moral exemplars, document how exemplars are open to, and

even seek, influence from their social environment that widens their circle of concern. For example, because of her early interests in justice and fairness, Virginia Durr found herself interacting with a variety of social justice activists whose own commitments influenced her to make broader and deeper commitments to social justice (Colby & Damon, 1992, chapter 5). This influence was likely based in both explicit argument with others and implicit influence by others on the way she saw and responded to injustice. These served as catalysts for evaluative and normative moral formation and thus character and personality change.

In addition to these interpersonal modes of influence on evaluative moral formation, the structure and social norms of the situations we inhabit can influence our personal development goals and plans. For instance, economies, architecture, and city planning all serve to distance many of us from the harm that meat-eating causes to those animals we consume (Bastian & Loughnan, 2016). Prejudices and stereotypes, but also architecture, serve to justify and maintain the existing social order, and vice versa (Jost & Banaji, 1994; Jost & Hunyady, 2005). Extensive work on the self in cultural context shows that who we become, and who we want to become, is culturally shaped, even in terms of how we critique or struggle against the prevailing culture (Gibbs et al., 2007).

9.4.3 Moral Formation as Cultural Conservation and Critique

Cultures vary systematically in terms of whether the relations between people are judged based upon independent experience and desire or interdependent social relations.⁴² “The nail that sticks out is likely to be hammered down in Japan whereas the squeaky wheel attracts grease and attention in the United States” (Markus & Kitayama, 2010, p. 420). The goals, plans, hopes, and narratives of members of a culture are shaped by these ideals. *Master narratives*, those narratives that are supported by a culture, can shape the personal narrative that an individual adopts. But they can also become the focus of conflict as people struggle to change or maintain those narratives. For instance, the heterosexual narrative of how family life develops has been challenged by the gay rights movement in the United States (Hammack, 2008), and the acceptable career narrative for women has changed in Western culture over the last decades (Keane, 2015). But in both of these instances, the narrative resources to critique the prevailing culture arose from within that culture in terms of language about human rights – but also religion (Carter, 1993, pp. 39–40). Thus, narratives about moral ideals are influenced by disagreements within a culture and those disagreements often draw on resources from that culture. Across a range of cultures, those who are more politically liberal tend to emphasize justice and care goals, while those who are more politically conservative tend to emphasize loyalty, authority, and sanctity in their judgments and ideals (Graham, Haidt, & Nosek, 2009; Haidt & Joseph, 2004). The difficulty of this meta-cultural disagreement in values lies in part with the inability of each side to articulate a narrative that appeals to the other side (Haidt, 2012). In these and many other ways, moral formation is socially and culturally embedded.

⁴²There is, of course, considerable variation *within* cultures in this dimension, and different aspects of a culture (e.g. legal vs. familial) may embed different assumptions about this. See Chapter 3 for more detail on these differences.

9.4.4 Moral Formation and Moral Luck

The civil rights activist Virginia Durr did not pay attention to social injustice in her youth but, influenced over the years by her social network and her own commitments to justice and honor, she became exquisitely sensitive to recognizing injustice and working against it (Colby & Damon, 1992). To be sure, some of this was consciously guided, but other episodes seemed automatic and well-practiced, such as her spontaneous reaction at a women's church meeting in Alabama when a black participant was asked to leave during lunch since the venue did not allow integrated eating – she walked out with the woman and ate lunch outside (Colby & Damon, 1992, p. 114).⁴³

Ms. Durr's career in civil rights was also shaped by the accidents of her biography: attending a racially integrated liberal arts college, being introduced in Washington DC to significant figures in the civil rights movement, etc. So the moral luck that accompanies us all can dramatically change our life stories in ways we do not anticipate.⁴⁴ We might find that we are led toward some (im)moral goals as events overtake us, or that our progress toward an (im)moral goal is restricted by events. Or one might not have any plans for becoming a particular kind of person, and instead simply find oneself drawn along a process. We have little systematic knowledge of the variety of paths in adulthood that lead one to moral action, or the role of intention in these paths. Or of the various ways that plans, goals, and narrative might reassert themselves in reaction to events. "We are never more (and sometimes much less) than the co-authors of our own narratives. Only in fantasy do we live what story we please" (MacIntyre, 1981, p. 199).⁴⁵ Moral formation is thus always multidimensional. It is sometimes more and sometimes less intentional, continuous, and controllable.

9.4.5 Critique from the Three Traditions

Monasticism is precisely about constructing a moral ecology for moral and spiritual development. The point of monastic community is to structure life and its influences in the service of the spiritual goal. In its long history, Christian monasticism has seen multiple waves of regression and reform (LeClercq, 1982) that might yield insight to the patterns that support institutions that are effective at moral formation and those that corrupt it. *Bildung*, in all its variations, is also about constructing an environment conducive to moral growth. There are only hints of this dense interdependence of moral ecology and moral formation in the moral psychology literature. It is most explicit in the work of Nick Epley and colleagues on ethics as a design problem (Epley & Kumar, 2019; Epley & Tannenbaum, 2017).⁴⁶ The complicated relationships between, for instance, organized religion, prejudice, and terrorism are well documented (see Paloutzian, 2017 for a review) but make no appearance in current

⁴³There is no evidence in the way she tells the story that her action was considered or weighed on its merits, she simply acted, "we had to go."

⁴⁴See Williams (1981) for philosophical reflections on moral luck.

⁴⁵My thanks to Anthony Rudd for his reference to this poignant quote in his essay *In Defense of Narrative* (Rudd, 2007).

⁴⁶See also Killen and Dahl (2021), who make a strong case for social reform based in moral reasoning, though they acknowledge the social and emotional aspects of SERE.

moral psychology. A fully developed moral psychology would see these issues as central. The complex histories of monasticism and *Bildung* and the existential revival of the Socratic maieutic could each provide fertile ground for theoretical efforts to develop this relationship.

9.5 Discussion

All human lives show moral formation, but not all lives are centered on morality; perhaps not even all highly moral lives are centered on morality.⁴⁷ But one can learn to be (more or less) moral, even in adulthood. This journey from self-centeredness to socially necessary moral commitment and possibly to moral exemplarity occurs under multiple constraints and influences and comes in many flavors. This text is in part a catalogue of these influences and processes, and of the myriad ways that personality, moral identity, skills and knowledge, and moral ecology interactively influence and are influenced by the processes of emotion, reason, and moral formation.

9.5.1 Conclusion

1. *We should replace a concern for foundationalist criteria and bias with a concern for the role of personal, interpersonal, and cultural critique.*

For moral formation, critique and struggle on the journey is a better metaphor than foundations for a building. The acceptance in moral psychology of what Walker (2004) calls formalist, or foundationalist, rational models (like those of Kohlberg etc.) has led to framing the issues of moral formation in terms of increasing commitment to a foundational criterion.⁴⁸ This approach is taken in part to escape from a simple relativism that would make us prisoners of socialization or instinct. For example, in Kohlberg, knowledge of and concern for justice is both the philosophical foundation and the developmental goal. Bias is then framed as coming from any source other than conscious reasoning about justice. Emotional, social, or cultural influence, religion, personality characteristics, etc. are all seen as bias (Walker, 2004). On the contrary, each of these can be a source of support for moral action, and reason can be a source of bias. In the vain pursuit of a foundationalist stance, we have cut ourselves off from the rich complexity of how and why we take moral action or progress in moral commitment.

An alternative, descriptive, model will involve the recognition of the complex of supports and constraints on moral life, the role of an active agent in navigating them, and the process of self-transcendence and adoption of moral goals. In every chapter in this text, we see the interweaving of these supports, constraints, and

⁴⁷One can imagine lives centered on spiritual or religious values that in the right circumstances result in extensive moral action, or lives devoted to ethical rule-following (rather than to abstract values) that might have a moral profile distinctly different from other moral exemplars (e.g. the exemplary FBI agent).

⁴⁸See Lapsley and Hardy (2017) and Lapsley (2020) for a parallel critique of moral development research as isolated within the “cognitive developmental approach to socialization.”

navigation. Central processes in this interweaving are personal, interpersonal, and cultural critique. These processes of critique are usually supported by a range of resources: values, master narratives, social influence, emotion, moral ecological influence, etc. These critiques can drive moral formation and we need to better understand the patterns in how they unfold. And it is these processes that allow us to continually critique our personal and societal ultimate values.

2. *We cannot avoid grappling with normative claims in studying moral formation.*

The research program of Frimer, Walker, and colleagues makes a compelling case that a central move in moral commitment is from agency alone to agency in the service of communion. Their empirical description of the movement leads them to say people *should* make this transition, and that this transition *should be valued as morally good* – clearly making a normative claim based on, or at least in tandem with, an empirical pattern.⁴⁹ Further, they claim that an unavoidable task in moral psychology will be grappling with how normative moral value is embedded in our empirical descriptions of moral development:

theories of moral personhood must endorse not only goals for how individuals should (deontologically) reason right from wrong but also what they should (teleologically) value as being good from bad. The new challenge becomes the highly controversial task of articulating and defending an account of which values are preferable in sufficiently broad terms to avoid ethnocentrism. To be clear, our point is that this task is *categorically unavoidable* for any theory of moral psychology ... [W]hich non-neutral conception of the good is most appropriate remains important conceptual work for the field. [emphasis in the original]⁵⁰ (Frimer & Walker, 2008, pp. 337–338)

We agree with the necessity of this task. It is, at the very least, the puzzle at the heart of the moral exemplar approach: studying exemplars necessarily assumes there is a way to describe excellence on a moral dimension and to identify exemplars on that dimension. The first dimension of the TOI model, self-transcendence, centers the problem of determining the highest good among the many values for which committed individuals might transcend the self. Adding the dimension of conservation-openness helps further open up the space of ultimate values and moral exemplars who might serve them. It is clear that moral inclusion/exclusion is a crucial indicator of for whom the self-transcendent commitment applies. Understanding self-transcendence as a process of transcending one's relative ends to serve a higher good will help to conceptualize the normative question as one of moral action, about how one brings the normative into action (see Section 9.5.1.4). Even rule-based morality involves selection of which rules are appropriate in what circumstances. It is difficult to imagine any adequate description of moral formation that does not address the puzzles associated with these three dimensions.

⁴⁹This *should* can be read as an instrumental should, that human development in the usual way involves this transition. But it has a flavor of a moral should, of the infinite. Colby and Damon (2015) make a similar claim that the field must grapple with the normative.

⁵⁰Note that here the process of evaluation is still limited to reason, and that the criterion seems to be the limited goal of “avoiding ethnocentrism.”

Choosing ways of evaluating these puzzles will be a central interdisciplinary task in moral psychology.⁵¹

3. *Moral formation likely occurs across multiple domains and on multiple dimensions.*

A significant obstacle in research on moral formation has been the desire to identify and map moral formation based upon a single domain or dimension. The fault for this cannot entirely be placed on what Lapsley and Hardy (2017) call the “cognitive developmental approach to socialization.” One can find it in almost all research programs in the area, even those not descended from the Kohlberg project. But complexity is unavoidable. Moral formation occurs with specific, concrete, commitments to a cause *and* at the abstract value level; the processes operate at implicit and explicit levels; there are multiple dimensions on which one can measure moral change and multiple aspects or levels of personality that can drive moral commitment. We have selected three dimensions among these (TOI: self-transcendence, openness, and inclusion) that we hope will provide the right level of abstraction and the necessary variation to stretch our understanding of moral formation.

4. *An adequate understanding of moral formation must include the dense weave of evaluative, normative, and educative aspects as they guide the processes of self-Transcendence.*

Moral formation is guided by an ongoing process of commitment to values, goals, or projects, and by the relations among these commitments. It is shaped by individual processes of evaluation of how well we are upholding the values or achieving the goals, and whether those goals are appropriate or even of ultimate concern. It is also embedded in a social matrix that schools us in those commitments and evaluations. Moral formation encompasses personality, reason, emotion, skills, moral ecology, and moral identity. It is an internal action of commitment, but one calling for expression, practice, and evaluation in moral action within specific domains of life and social contexts. Recognizing which aspects of moral formation are occurring at what times and how they interweave will be crucial to understanding individuals’ moral life trajectories as they navigate their lives. Understanding this complexity in moral formation will require interdisciplinary rigor in investigation, rather than simple hand waving. Our critiques of the processes at the end of each section suggest fertile ground for this interdisciplinary investigation.

9.5.2 Application

We met the Munson family in the Introduction and we return to them here. They were seemingly living the good life of the upper middle class when they decided to take in foster children. Their religious backgrounds suggest that communal value commitments played a part in this decision. But it may have been a more self-centered desire for a family as fulfillment of a community role. It is also possible that their religious commitments were their first priority and that these in turn lead to commitment that

⁵¹This task is the door to philosophy and theology. If the task of choosing “which values are preferable” is indeed unavoidable, then this door is also unavoidable. Theologians can do this without necessarily being sectarian, as exemplified by Herdt’s (2019) expansive conclusion to her book on *Bildung*.

a psychologist might describe as communal (though actually rooted in a religious value dimension). The monk Thomas Merton's initial commitments seemed focused on personal religious need in a close-knit monastic community, but the social upheaval of his time, and the religious narratives in which he was embedded, offered him an opportunity to channel this into a profound concern for the oppressed. Thus, his journey of moral formation, as well as the Munsons', is shaped and supported by the variety of contexts, influences, and processes we outline in this book.

The Munson narrative is one of a slow, stepwise increase in commitment. First, they committed to some particular children, then, faced with physically disabled children, they redoubled their commitment to the children but also extended it to other children like them who were having trouble finding services to help them. Their work in this area led to volunteering at schools, serving on boards, and eventually founding a charitable organization to provide the needed services to children in the area.

Surely the levels of commitment and the actual goal of their actions changed over time. But if we were to interview them, it seems unlikely they would describe the process as one of continuous rational evaluation of their commitments and planned increases in action based on the outcome of those reflections. However, they might well agree with a narrative of a journey in which they were drawn to do more of the good than they had started out with. This deepening of what came to be an ultimate concern is part and parcel of moral formation approaches across contemplative and religious traditions, we can find it in concepts of existential appropriation as well as in the education of the heart. This deepening of what concerns us ultimately, the moral good we consider our highest, comes with the task of evaluating and re-evaluating not only our normative value system, but also the nature of our attachments to each of them. It is as much shaped by the nature of our evaluations, choices, and commitments as it is by our situation and culture. At each turn in the journey, the Munsons might have decided they had done enough. Then they would not have been confronted with the next situation that also had its calls for commitment or rejection.

9.5.3 Open Questions

The central task and puzzle of this chapter is to assemble the various aspects of moral formation in a way that is faithful to the wide variety of (im)moral trajectories. Only then can we begin carefully to think about how to empirically describe them and how to theoretically explain them. We suggest here some open questions that this attempt offers:

1. *Are there common psychological aspects to the commitments to the many different types of self-transcending values and ultimate concerns?*

We have noted that most moral exemplars do not start out with a conscious commitment to "being good" but begin instead with specific commitments as the Munsons or Thomas Merton do in our application case. At this more concrete level, there may be thousands of ways to get on the path to moral formation. In fact, an existential viewpoint would suggest that we are already and always on that path, whether we want it or not. Still, at some higher abstract level this multitude may fit the self-enhancement to self-transcendence description. We have regularly remarked that some things that are taken to be ultimate can lead to terrorism,

genocide, and other evils. Are the processes of commitment, evaluation, and schooling the same in all cases, good and evil? Are there any of these processes that might help us in distinguishing good vs. evil vs. indifferent commitments? What aspects of these processes and commitments constitute what is commonly called virtue? How are they distinguished from vice? And what might we learn from studying the negative trajectory of someone's moral formation?

2. *What is the relation of skill development to taking moral action?*

The study by Frimer et al. (2012) provides us with a way to frame this question. In their study of exemplars taken from the front pages of *Time Magazine*, they identify "agentic" and communal exemplars. Agentic exemplars are known for taking effective action, but not really for taking effective *moral* action. Do these two groups share skill sets, or are there some skills that are more associated with taking effective *moral* action? What might these moral skills be? Does developing these moral skills influence moral formation or are these two different tasks? The monastic tradition we covered early in the chapter highlights some form of humility as crucial in moral formation. There are empirical literatures we review in this chapter that identify the skills associated with a robust humility.

3. *How might the traditional accounts of moral formation help us understand and explain psychological processes of moral action?*

Most psychological analyses of moral formation tend to fragment the processes they study, and scientific psychology is often drawn to those aspects that are easiest to measure or seem most objective. They concentrate on cognition, or emotional regulation, and there is little within them that compels a theoretical change of perspective, broadening of focus, or deepening of understanding. Nor do they grapple with the existential depth of the struggle with the human condition and the infinite. This is how we end up with a pile of pearls and no string to bind them together. We have presented several more holistic approaches to moral formation that are non-scientific but nevertheless deeply empirically oriented because they are based in long traditions of observation and critique of human moral action. These traditions see moral formation as both the self's careful navigation of the moral world and an intentional reflection on and deepening of our commitments; as a longing search for wholeness and a willingness to let oneself and one's goals be challenged by the social and normative world one encounters. If they are complicated and paradoxical, it is likely because the phenomenon they have been closely observing and trying to guide is itself complicated, paradoxical, and full of variation.

We are not proposing these traditions as alternative psychological theories. Rather, each proposes its approach as a practical "school" for formation that suggests opportunities for psychological research. See our comments at the end of each section for specific critiques that draw from this well of practical, empirical experience. Rigorous interdisciplinary inquiry may help identify ways that existing theoretical and empirical approaches in psychology are overlooking or ignoring important aspects of the broader process of moral formation.

4. *How does moral ecology interact with moral formation?*

A theme throughout this text has been the applied question of how we can construct moral ecologies that support moral action. We now have some language to phrase that question in a way that may advance the conversation: How can we

construct *moral ecologies* that support effective *individual, social, and cultural critique* in a way that helps achieve appropriate *SERE* in moral commitments and action? There are many questions remaining within each of the terms in that sentence. And modifiers like “appropriate” are invitations – perhaps even demands – for interdisciplinary collaboration. Nevertheless, we hope this phrasing of the question opens the doors to conversations in service of moral action.

5. *How is a SERE achieved and maintained in the face of ultimate commitments?*

In several other chapters (e.g. Chapters 3, 7, and 8) we mention the processes of social and cultural critique. This chapter has primarily focused on processes of personal critique (as does Chapter 5). And we have argued often that seeking a *SERE* is a useful description for these processes. This balance occurs across multiple dimensions (social, emotional, cognitive). We do not propose a model for how this happens. We are simply specifying the things one must bring into equilibrium. The phrase reflective equilibrium is taken from philosophy (Rawls, 1971/1999) and suggests a balance among considerations. But clearly, if one is considering ultimate concerns, equally weighting all the concerns is out of the question. How does one find a weighted balance in the face of the ultimate? So, perhaps it is achieving the appropriate tension instead. Maintaining this tension among our values, goals, or projects might be what allows the process of recognizing disfluency (see Chapter 6) when the tension among our moral commitments becomes too great. This recognition of disfluency helps one to break out of automatic processes and begin conscious, controlled reflection on why things are not working. This reflection may mean re-evaluating a host of things, including one’s ultimate commitment (see the Chapter 5 for the complexity involved here). Thus, reflective equilibrium might be better expressed as the feeling of peace or clarity one achieves after the weighing and recalibration process.

9.6 Further Readings

These suggested readings are designed to lead the reader further into the literature that forms the main themes of this chapter. They combine some classic pieces and recent work. Complete citations are provided in the references section.

- McAdams (2013). “The psychological self as actor, agent, and author.” An understanding of change in the self at three different levels.
- Bluck et al. (2005). “A tale of three functions: The self-reported uses of autobiographical memory.” An empirical study to explore the uses of autobiographical narrative in both change and maintenance of the self.
- Colby and Damon (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. A classic study of moral exemplars in social service in the United States that includes discussion of change in moral commitments over time.
- Applebaum (2012). *Iron Curtain: The Crushing of Eastern Europe 1944–1956*. A study of Soviet repression in East Germany, Hungary, and Poland. It shows the complexities of changing moral ecologies and their influence on moral commitment and ultimate value.

- Gelfand (2021). “Cultural evolutionary mismatches in response to collective threat.” An overview of the idea of cultural tightness: a concept containing both strictness of norm following and narrowness of in-group. This varies across both cultures and over time.
- Herdt (2019). *Forming Humanity: Redeeming the German Bildung Tradition*. A comprehensive, complex, and subtle overview of the *Bildung* tradition from its classical sources to its integration in current theological and philosophical thought.
- Walker and Frimer (2015). “Developmental trajectories of agency and communion in moral motivation.” Theoretical review of their research program that documents change from agentic commitments to agency in service of communion in adolescents, young adults, and moral exemplars.
- Wrzus and Roberts (2017). “Processes of personality development in adulthood: The TESSERA framework.” A complex model of personality change that incorporates both implicit and intentional change.

References

- Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, 78(1), 53–63.
- Ackerman, J. M., & Bargh, J. A. (2010). The purpose driven life: Commentary on Kenrick et al. *Perspectives on Psychological Science*, 5(3), 323–326. doi:10.1177/1745691610369472
- Addison, H. (2000). *Show Me Your Way: The Complete Guide to Exploring Interfaith Spiritual Direction*. Woodstock, VT: Skylight Paths Publishing.
- Adler, J. M. (2013). Living into the story: Agency and coherence in a longitudinal study of narrative identity development and mental health over the course of psychotherapy. *Journal of Personality and Social Psychology*, 102(2), 367–389.
- Adler, J. M., Lodi-Smith, J., Phillipe, F. L. et al., (2015). The incremental validity of narrative identity in predicting well-being: A review of the field and recommendations for the future. *Personality and Social Psychology Review*, 20(2), 142–175. doi:10.1177/1088868315585068
- Adler, J. M., Wagner, J. W., & McAdams, D. P. (2007). Personality and the coherence of psychotherapy narratives. *Journal of Research in Personality*, 41(6), 1179–1198. doi:10.1016/j.jrp.2007.02.006
- Allen, S., & Fry, L. (2022). A framework for leader, spiritual, and moral development. *Journal of Business Ethics*, 184(3), 649–663. doi:10.1007/s10551-022-05116-y
- Appiah, K. A. (2001). Cosmopolitan reading. In V. Dharwadker (Ed.), *Cosmopolitan Geographies: New Locations in Literature and Culture* (pp. 223–240). New York: Routledge.
- Applebaum, A. (2012). *Iron Curtain: The Crushing of Eastern Europe 1944–1956*. New York: Doubleday.
- Aquino, K., Reed, A., Thau, S. et al., (2006). A grotesque and dark beauty: How moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. *Journal of Experimental Social Psychology*, 43(3), 385–392. doi:10.1016/j.jesp.2006.05.013
- Augustine, M. J. (1986). Zen and Benedictine monks as mythopoeic models of nonegocentered worldviews and lifestyles. *Buddhist-Christian Studies*, 23–49.
- Balboni, M. J., Bandini, J., Mitchell, C. et al., (2015). Religion, spirituality, and the hidden curriculum: Medical student and faculty reflections. *Journal of Pain and Symptom Management*, 50(4), 507–515. doi:10.1016/j.jpainsymman.2015.04.020

- Bastian, B., & Loughnan, S. (2016). Resolving the meat-paradox: A motivational account of morally troublesome behavior and its maintenance. *Personality and Social Psychology Review: Journal of the Society for Personality and Social Psychology, Inc.* doi:10.1177/1088868316647562
- Baumeister, R. F., Masicampo, E. J., & Vohs, K. D. (2011). Do conscious thoughts cause behavior? *Annual Review of Psychology*, 62(1), 331–361. doi:10.1146/annurev.psych.093008.131126
- Bluck, S., Alea, N., Habermas, T. et al., (2005). A tale of three functions: The self-reported uses of autobiographical memory. *Social Cognition*, 23(1), 91–117.
- Borg, I., Bardi, A., & Schwartz, S. H. (2017). Does the value circle exist within persons or only across persons? *Journal of Personality*, 85(2) doi:10.1111/jopy.12228
- Bosi, L. (2016). Social movements and interrelated effects: The process of social change and the post-movement lives of provisional IRA volunteers. *Revista Internacional de Sociología*, 74(4). doi:10.3989/ris.2016.74.4.047
- Cantor, N., Norem, J. K., Niedenthal, P. M. et al., (1987). Life tasks, self-concept ideals, and cognitive strategies in a life transition. *Journal of Personality and Social Psychology*, 53(6), 1178–1191.
- Carter, S. (1993). *The Culture of Disbelief*. New York: Basic Books.
- Carver, C. S., & Scheier, M., F. (2002). Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review*, 6(4), 304–315.
- Casey, M. (2005). Models of monastic formation. In M. Casey (Ed.), *An Unexciting Life: Reflections of Benedictine Spirituality* (pp. 365–405). Petersham, MA: St. Bede's Publications.
- Cervone, D. (2004). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, 56, 423–452.
- Cervone, D. (2021). Five paths to personality coherence: Integrative implications of the knowledge-and-appraisal personality architecture. *European Journal of Personality*, 36(3), 319–346. doi:10.1177/08902070211015599
- Cervone, D., & Tripathi, R. (2009). The moral functioning of the person as a whole: On moral psychology and personality science. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 30–51). New York: Cambridge University Press.
- Cieciuch, J., Schwartz, S. H., & Davidov, E. (2015). The social Psychology of values. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., Vol 25, pp. 41–46). Oxford: Elsevier. doi:10.1016/b978-0-08-097086-8.25098-8
- Clayton, S., & Opatow, S. (2003). Justice and identity: Changing perspectives on what is fair. *Personality and Social Psychology Review*, 7(4), 298–310. doi:10.1207/S15327957PSPR0704_03
- Cline, E. (2018). *A World on Fire: Sharing the Ignatian Spiritual Exercises*. Washington, DC: Catholic University Press.
- Cohen-Chen, S., Pliskin, R., & Goldenberg, A. (2020). Feel good or do good? A valence-function framework for understanding emotions. *Current Directions in Psychological Science*, 29(4), 388–393. doi:10.1177/0963721420924770
- Cojuharenco, I., Cornelissen, G., & Karelaia, N. (2016). Yes, I can: Feeling connected to others increases perceived effectiveness and socially responsible behavior. *Journal of Environmental Psychology*, 48, 75–86. doi:10.1016/j.jenvp.2016.09.002
- Colby, A., & Damon, W. (1992). *Some Do Care: Contemporary Lives of Moral Commitment*. New York: Free Press.
- Colby, A., & Damon, W. (2015). *The Power of Ideals: The Real Story of Moral Choice*. Oxford, UK: Oxford.
- Côté, S., Decelles, K. A., McCarthy, J. M. et al., (2011). The Jekyll and Hyde of emotional intelligence: Emotion-regulation knowledge facilitates both prosocial and interpersonally deviant behavior. *Psychological Science*, 22(8), 1073–1080. doi:10.1177/0956797611416251

- Crimston, D., Bain, P. G., Hornsey, M. J. et al., (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology*, 111(4), 636–653. doi:10.1037/pspp0000086
- Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology*, 78(4), 791–808. doi:10.1037/0022-3514.78.4.791
- Day, L. C., & Impett, E. A. (2018). Giving when it costs: How interdependent self-construal shapes willingness to sacrifice and satisfaction with sacrifice in romantic relationships. *Journal of Social and Personal Relationships*, 35(5), 722–742. doi:10.1177/0265407517694965
- Depow, G. J., Frances, Z., & Inzlicht, M. (2021). The experience of empathy in everyday life. *Psychological Science*, 32(8), 1198–1213. doi:10.1177/0956797621995202
- Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology*, 61, 467–490. doi:10.1146/annurev.psych.093008.100445
- Dunlop, W. L., Walker, L. J., & Matsuba, M. K. (2013). The development of moral motivation across the adult lifespan. *European Journal of Developmental Psychology*, 10(2), 285–300. doi:10.1080/17405629.2012.746205
- Dyckman, K., Garvin, M., & Liebert, E. (2001). *The Spiritual Exercises Reclaimed: Uncovering Liberating Possibilities for Women*. Mahwah, NJ: Paulist Press.
- Einolf, C. J. (2010). Does extensivity form part of the altruistic personality? An empirical test of Oliner and Oliner's theory. *Social Science Research*, 39(1), 142–151. doi:10.1016/j.ssresearch.2009.02.003
- Emmons, R. A. (1986). Personal strivings: A approach to personality and subjective well being. *Journal of Personality and Social Psychology*, 51(5), 1058–1068.
- Ent, M. R., Baumeister, R. F., & Tice, D. M. (2015). Trait self-control and the avoidance of temptation. *Personality and Individual Differences*, 74, 12–15. doi:10.1016/j.paid.2014.09.031
- Epley, N., & Kumar, A. (2019). How to design an ethical culture. *Harvard Business Review*, 3(May–June), 144–150.
- Epley, N., & Tannenbaum, D. (2017). Treating ethics as a design problem. *Behavioral Science & Policy*, 3(2), 73–84.
- Eskreis-Winkler, L., & Fishbach, A. (2022). You think failure is hard? So is learning from it. *Perspectives on Psychological Science*, 17(6), 1511–1524. doi:10.1177/17456916211059817
- Fiske, A. P., & Rai, T. S. (2014). Violence for goodness' sake. *New Scientist*, 224(2997), 30–31. doi:10.1016/s0262-4079(14)62289-2
- Fowler, Z., Law, K. F., & Gaesser, B. (2021). Against empathy bias: The moral value of equitable empathy. *Psychological Science*, 32(5), 766–779. doi:10.1177/0956797620979965
- Frankl. (1997). *Man's Search for Ultimate Meaning*. New York: Insight Books.
- Frimer, J. A. (2012). Reverse engineering good people. In *New Directions for Child and Adolescent Development*. Unpublished paper. University of British Columbia. [http://jeremyfrimer.com/Welcome_files/Frimer%20\(in%20press%3B%20NDCAD\).pdf](http://jeremyfrimer.com/Welcome_files/Frimer%20(in%20press%3B%20NDCAD).pdf)
- Frimer, J. A., Tell, C. E., & Haidt, J. (2015). Liberals condemn sacrilege too: The harmless desecration of Cerro Torre. *Social Psychological and Personality Science*, 6(8), 878–886. doi:10.1177/1948550615597974
- Frimer, J. A., & Walker, L. J. (2008). Towards a new paradigm of moral personhood. *Journal of Moral Education. Special Issue: Towards an Integrated Model of Moral Functioning*, 37(3), 333–356. doi:10.1080/03057240802227494
- Frimer, J. A., & Walker, L. J. (2009). Reconciling the self and morality: an empirical model of moral centrality development. *Developmental Psychology*, 45(6), 1669–1681. doi:10.1037/a0017418
- Frimer, J. A., Walker, L. J., Dunlop, W. L. et al., (2011). The integration of agency and communion in moral personality: Evidence of enlightened self-interest. *Journal of Personality and Social Psychology*, 101(1), 149–163.

- Frimer, J. A., Walker, L. J., Lee, B. H. et al., (2012). Hierarchical integration of agency and communion: a study of influential moral figures. *Journal of Personality*, 80(4), 1117–1145. doi:10.1111/j.1467-6494.2012.00764.x
- Furchert, A. (2005). From evidence to existence. Kierkegaard read as existential corrective to a new objectivism in psychiatry. Paper presented at the 8th International Conference on Philosophy, Psychiatry, and Psychology, Yale, New Haven CT.
- Furchert, A. (2019). Hildegard of Bingen. In C. Talliaferro (Ed.), *Wiley-Blackwell Encyclopedia of Philosophy of Religion*. New York: Wiley-Blackwell.
- Furchert, A. (2022). Bringing God into existence. Kierkegaard's philosophy of appropriation revisited. Retrieved from https://www.academia.edu/55140937/Bringing_God_into_Existence_Kierkegaards_philosophy_of_appropriation_revisited
- Gadamer, H.-G. (1975). *Wahrheit und Methode. Grundzuge einer philosophischen Hermenutik*. Tübingen: J. C. B. Mohr.
- Gadamer, H.-G. (1996). *The Enigma of Health: The Art of Healing in a Scientific Age*. Stanford, CA: Stanford University Press.
- Gelfand, M. (2012). Culture's constraints: International differences in the strength of social norms. *Current Directions in Psychological Science*, 21(6), 420–424. doi:10.1177/0963721412460048
- Gelfand, M. (2021). Cultural evolutionary mismatches in response to collective threat. *Current Directions in Psychological Science*, 30(5), 401–409. doi:10.1177/09637214211025032
- Gibbs, J. C., Basinger, K. S., Grime, R. L. et al., (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review*, 27(4), 443–500. doi:10.1016/j.dr.2007.04.001
- Gilbert, D. T., & Wilson, T. D. (2007). Prospection: experiencing the future. *Science*, 317(5843), 1351–1354. doi:10.1126/science.1144161
- Giugni, M. G. (2004). Personal and biographical consequences. In D. A. Snow, S. A. Soule, & H. Kreisi (Eds.), *The Blackwell Companion to Social Movements* (pp. 489–507). Oxford: Blackwell.
- Goethe, J. W. (1795/2007). *Wilhelm Meisters Lehrjahre*. Berlin: Suhrkamp.
- Golban, P. (2018). *A History of the Bildungsroman: From Ancient Beginnings to Romanticism*. Newcastle upon Tyne: Cambridge Scholars.
- Gollwitzer, P. M. (2015). Setting one's mind on action: Planning out goal striving in advance. In R. Scott & S. Kosslyn, *Trends in the Social and Behavioral Sciences* (pp. 1–14). Palto, CA: Center for Advanced Study in the Behavioral Sciences, Stanford University. <http://emerging-trends.stanford.edu/files/original/28cddfd4aa523e572556b73924e41c6d4dd7b9e9.pdf>
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029–1046.
- Graham, J., Waytz, A., Meindl, P. et al., (2017). Centripetal and centrifugal forces in the moral circle: Competing constraints on moral learning. *Cognition*, 167, 58–65. doi:10.1016/j.cognition.2016.12.001
- Grube, J. A., & Piliavin, J. A. (2000). Role identity, organizational experiences, and volunteer performance. *Personality and Social Psychology Bulletin*, 26(9), 1109–1119.
- Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided By Politics and Religion*. New York: Pantheon.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.
- Hammack, P. L. (2008). Narrative and the cultural psychology of identity. *Personality and Social Psychology Review*, 12(3), 222–247. doi:10.1177/1088868308316892
- Harris, C. E. (2008). The good engineer: Giving virtue its due in engineering ethics. *Science and Engineering Ethics*, 14(2), 153–164.
- Hart, D. (2005a). Adding identity to the moral domain. *Human Development*, 48(4), 257–261.

- Hart, D. (2005b). The development of moral identity. In C. Pope-Edwards & G. Carlo (Eds.), *Nebraska Symposium on Motivation, Vol 51: Moral Motivation Through the Life Span* (pp. 166–196). Lincoln, NE: University of Nebraska Press.
- Hart, D., & Carlo, G. (2005). Moral development in adolescence. *Journal of Research on Adolescence. Special Issue: Moral Development*, 15(3), 223–233. doi:10.1111/j.1532-7795.2005.00094.x
- Hartshorne, H., & May, M. A. (1928). *Studies in the Nature of Character: Studies in Deceit, Vol. 1*. New York: Macmillan.
- Herdt, J. A. (2019). *Forming Humanity: Redeeming the German Bildung Tradition*. Chicago, IL: University of Chicago.
- Higgins, E. T., Tykocinski, O., & Vookles, J. (1990). Patterns of self-beliefs: The psychological significance of relations among the actual, ideal, ought, can, and future selves. In J. M. Olsen & M. P. Zanna (Eds.), *Self-Inference Processes: The Ontario Symposium, Vol. 6* (pp. 153–190). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Higgins-D'Alessandro, A. (2015). Lawrence Kohlberg's legacy: Radicalizing the educational mainstream. In B. Zizack, D. Garz, & E. Nowak (Eds.), *Kolberg Revisited* (pp. 27–49). Rotterdam: Sense.
- Houdek, F. (1996). *Guided by the Spirit: A Jesuit Perspective on Spiritual Direction*. Chicago, IL: Loyola Press.
- Huff, C., & Barnard, L. K. (2009). Good computing: moral exemplars in the computing profession. *IEEE Technology and Society Magazine* (Fall), 47–54.
- Huff, C., & Furchert, A. (2014). Toward a pedagogy of ethical practice: The application problem. *Communications of the ACM*, 27(7), 25–27.
- Huff, C., & Furchert, A. (2023). Existential Appropriation in a digital age: The heart at the heart of digital ethics. In S. Rogerson (Ed.), *Handbook of Digital Social Science*. Cheltenham: Edgar Elgar.
- Humboldt, W. V. (1792/1960). Theorie der Bildung des Menschen. In A. Flitner & K. Giel (Eds.), *Wilhelm von Humboldt: Werke in Fünf Banden* (pp. 234–240). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Hutchins, R. (1953). *The University of Utopia*. Chicago, IL: University of Chicago.
- Jackson, P. (1968). *Life in Classrooms*. New York: Holt, Rinehart and Winston.
- JafariNaimi, N., Nathan, L., & Hargraves, I. (2015). Values as hypotheses: Design, inquiry, and the service of values. *Design Issues*, 31(4), 91–104. doi:10.1162/DESI_a_00354
- James, W. (1892/1963). *Psychology*. Greenwich, CT: Fawcett.
- Janoff-Bulman, R., & Carnes, N. C. (2018). The model of moral motives: A map of the moral domain. In K. Gray & J. Graham (Eds.), *Atlas of Moral Psychology* (pp. 223–232). New York: Guilford Press.
- Jaspers, K. (1938). *Existenzphilosophie*. Berlin: de Gruyter.
- 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), 1–27. doi:10.1111/j.2044-8309.1994.tb01008.x
- Jost, J. T., & Hunyady, O. (2005). Antecedents and consequences of system-justifying ideologies. *Current Directions in Psychological Science*, 14(5), 260–265. doi:10.1111/j.0963-7214.2005.00377.x
- Karl, J. (2022 August 21). TRANSCRIPT: ABC News' Jonathon Karl interviews Rep. Liz Cheney. ABC News.
- Keane, W. (2015). *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton University Press.
- Kenrick, D. T., Griskevicius, V., Neuberg, S. L. et al., (2010). Renovating the pyramid of needs: Contemporary extensions built upon ancient foundations. *Perspectives on Psychological Science*, 5(3), 292–314. doi:10.1177/1745691610369469

- Kierkegaard, S. (1846/1992). *Concluding Unscientific Postscript to Philosophical Fragments: A Mimetic-Pathetic-Dialectic Compilation, an Existential Plea, by Johannes Climacus* (H. V. Hong & E. H. Hong, Trans.). Princeton, NJ: Princeton University Press.
- Killen, M., & Dahl, A. (2021). Moral reasoning enables developmental and social change. *Perspectives on Psychological Science*, 16(6), 1209–1225. doi:10.1177/1745691620964076
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive Development and Epistemology* (pp. 151–235). New York: Academic Press.
- Koltko-Rivera, M. E. (2006). Rediscovering the later version of Maslow's hierarchy of needs: Self-transcendence and opportunities for theory, research, and unification. *Review of General Psychology*, 10(4), 302–317. doi:10.1037/1089-2680.10.4.302
- Kuhl, J., & Koole, S. L. (2004). Workings of the will: A functional approach. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of Experimental Existential Psychology* (pp. 411–430). New York: Guilford Press.
- Kunst, J. R., Boos, B., Kimel, S. Y. et al. (2018). Engaging in extreme activism in support of others' political struggles: The role of politically motivated fusion with out-groups. *Plos One*, 13(1), e0190639. doi:10.1371/journal.pone.0190639
- Lapsley, D. K. (2016). Moral self-identity and the social-cognitive theory of virtue. In J. Annas, D. Narvaez, & N. E. Snow (Eds.), *Developing the Virtues: Integrating Perspectives*. New York: Oxford University Press.
- Lapsley, D. K. (2020). Moral formation in the family: A research agenda in time future. In D. J. Laible, G. Carlo, & L. M. Padilla-Walker (Eds.), *Oxford Handbook of Parenting and Moral Development* (pp. 393–418). New York: Oxford University Press.
- Lapsley, D. K., & Hardy, S. A. (2017). Identity formation and moral development in emerging adulthood. In L. M. Padilla-Walker & L. J. Nelson (Eds.), *Flourishing in Emerging Adulthood: Positive Development During the Third Decade of Life* (pp. 14–39). New York: Oxford University Press.
- Lawrence, C., Mhlaba, T., Stewart, K. A. et al., (2018). The hidden curricula of medical education: A scoping review. *Academic Medicine*, 93(4), 648–656. doi:10.1097/ACM.0000000000002004
- LeClercq, J. (1982). *The Love of Learning and the Desire for God*. New York: Fordham University Press.
- Little, B. R. (2008). Personal projects and free traits: Personality and motivation reconsidered. *Social and Personality Psychology Compass*, 2(3), 1235–1254.
- MacIntyre, A. C. (1981). *After Virtue: A Study in Moral Theory* (Vol. 1985). London: Duckworth.
- Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, 5(4), 420–430. doi:10.1177/1745691610375557
- Maslow, A. H. (1954). *Motivation and Personality*. New York: Addison-Wesley.
- Maslow, A. H. (1969). The farther reaches of human nature. *The Journal of Transpersonal Psychology*, 1(1), 1–9.
- McAdam, D. (1989). The biographical consequences of activism. *American Sociological Review*, 54, 744–760.
- McAdams, D. P. (2013). The psychological self as actor, agent, and author. *Perspectives on Psychological Science*, 8(3), 272–295. doi:10.1177/1745691612464657
- McAdams, D. P. (2015). *The Art and Science of Personality Development*. New York: Guilford Press.
- McAdams, D. P., & Pals, J. L. (2006). A new big five: Fundamental principles for an integrative science of personality. *American Psychologist*, 61(3), 204–217.
- McAdams, D. P., Reynolds, J., Lewis, M. et al., (2001). When bad things turn good and good things turn bad: Sequences of redemption and contamination in life narrative and their

- relation to psychosocial adaptation in midlife adults and in students. *Personality and Social Psychology Bulletin*, 27(4), 474–485. doi:10.1177/0146167201274008
- McClelland, D. C. (1987). *Human Motivation*. Cambridge: Cambridge University Press.
- Merton, T. (1948/1998). *The Seven Story Mountain*. New York: Houghton Mifflin Harcourt.
- Merton, T. (1964). *Seeds of Destruction*. New York: Farrar, Straus and Giroux.
- Merton, T. (1966). *Conjectures of a Guilty Bystander*. New York: Crown.
- Merton, T. (1975). *The Asian Journal of Thomas Merton* (P. Hart & J. Laughlin Eds.). New York: New Directions.
- Merton, T. (2009). *The Rule of Saint Benedict: Initiation into the Monastic Tradition 4*. Collegeville MN: Liturgical Press.
- Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1–22.
- Murray, H. A. (1938). *Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age*. Oxford: Oxford University Press.
- Narvaez, D. (2010). Moral complexity: The fatal attraction of truthiness and the importance of mature moral functioning. *Perspectives on Psychological Science*, 5(2), 163–181.
- Narvaez, D., Four Arrows, Halton, E., Collier, B. et al., (Eds.). (2019). *Indigenous Sustainable Wisdom: First-Nation Know-How for Global Flourishing*. New York: Peter Lang.
- Nielsen, K. S., & Bauer, J. M. (2018). The merits of goal support as a self-control strategy. *Social Psychological and Personality Science*, 10(5), 671–680.
- Nucci, L. P. (2004). Reflections on the moral self construct. In D. K. Lapsley & D. Narvaez (Eds.), *Moral Development, Self and Identity* (pp. 111–132). Mahwah, NJ: Erlbaum.
- Nussbaum, M. (1994). Patriotism and cosmopolitanism. *Boston Review* (October 1). <https://www.bostonreview.net/articles/martha-nussbaum-patriotism-and-cosmopolitanism/>
- Oliner, S. P., & Oliner, P. M. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press
- Opatow, S. (2001). Social injustice. In D. J. Christie, R. V. Wagner, & D. A. Winter (Eds.), *Peace, Conflict, and Violence: Peace Psychology for the 21st Century*. Englewood Cliffs, NJ: Prentice-Hall.
- Opatow, S. (2005). Moral exclusion to moral inclusion: Theory for teaching peace. *Theory into Practice*, 44(4), 303–318.
- Paloutzian, R. F. (2017). *Invitation to the Psychology of Religion* (3rd ed.). New York: Guilford Press.
- Peterson, C., & Park, N. (2010). What happened to self-actualization: Commentary on Kenrick et al. (2010). *Perspectives on Psychological Science*, 5(3), 320–322. doi:10.1177/1745691610369471
- Piliavin, J. A., & Callero, P. L. (1991). *Giving Blood: The Development of an Altruistic Identity*. Baltimore, MD: John Hopkins University Press.
- Plante, T. (2021). Using the Examen, a Jesuit prayer, in spiritually integrated and secular psychotherapy. *Pastoral Psychology*, 71, 119–125. doi:10.1007/s11089-021-00967-0
- Pratt, M. W., Hunsberger, B., Pancer, S. M. et al., (2003). A longitudinal analysis of personal values socialization: Correlates of a moral self-ideal in late adolescence. *Social Development*, 12(4), 563–585.
- Rambachan, A. (2006). *The Advaita Worldview: God, World, and Humanity*. New York: State University of New York Press.
- Rasmussen, A. S., & Habermas, T. (2011). Factor structure of overall autobiographical memory usage: The directive, self, and social functions revisited. *Memory*, 19(6), 597–605.
- Rawls, J. (1971/1999). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Reed, A., & Aquino, K. (2003). Moral identity and the expanding circle of moral regard towards out-groups. *Journal of Personality and Social Psychology*, 84(6), 1270–1286.

- Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of Personality Development* (pp. 11–39). Mahwah NJ: Lawrence Erlbaum Associates.
- Rudd, A. (2007). In defence of narrative. *European Journal of Philosophy*, 17(1), 60–75.
- Ruffing, J. (2002). Recent literature and emerging issues in the Ministry of Spiritual Direction. *Spiritus: A Journal of Christian Spirituality*, 2(1), 99–107. doi:10.1353/scs.2002.0019
- Rule, J. T., & Bebeau, M. J. (2005). *Dentists Who Care: Inspiring Stories of Professional Commitment*. Chicago, IL: Quintessence Publishing Co.
- Rusbult, C. E., Finkel, E. J., & Kumashiro, M. (2009). The Michelangelo phenomenon. *Current Directions in Psychological Science*, 18(9), 305–309.
- Ryan, R. M., & Deci, E. L. (2003). On assimilating identities to the self: A self-determination theory perspective on internalization and integrity within cultures. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of Self and Identity* (pp. 253–274). New York: Guilford Press.
- Schachter, E., & Ben Hur, A. (2019). The varieties of religious significance: An idiographic approach to study religion's role in adolescent development. *Journal of Research in Adolescence*, 29(2), 291–307. doi:10.1111/jora.12479
- Schnitker, S. A., King, P. E., & Houlberg, B. (2019). Religion, spirituality, and thriving: transcendent narrative, virtue, and telos. *Journal of Research on Adolescence*, 29(2), 276–290. doi:10.1111/jora.12443
- Schwartz, S. H. (2006). Basic human values: Theory, measurement and applications. *Revue Française de Sociologie*, 4(47), 929–968.
- Schwartz, S. H. (2016). Basic individual values: Sources and consequences. In D. Sander & T. Brisch (Eds.), *Handbook of Value*. Oxford: Oxford University Press.
- Schwartz, S. H., Cieciuch, J., Vecchione, M. et al., (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, 103(4), 663–688.
- Seligman, M. E. P., Railton, P., Baumeister, R. F. et al., (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science*, 8(2), 119–141. doi:10.1177/1745691612474317
- Shapiro, S., & Weisbaum, E. (2020). History of mindfulness and psychology. In O. Braddick (Ed.), *Oxford Research Encyclopedia of Psychology*. Oxford: Oxford University Press.
- Shoda, Y., Lee-Tiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and consistency from intra- and inter-personal interactions. *Personality and Social Psychology Review*, 6(4), 316–325.
- Shweder, R. A. (2012). Relativism and universalism. In D. Fassin (Ed.), *A Companion to Moral Anthropology* (pp. 85–102). New York: Wiley-Blackwell.
- Shweder, R. A., & Menon, U. (2014). Old questions for the new anthropology of morality: a commentary. *Anthropological Theory*, 14(3), 356–370.
- Shweder, R. A., & Power, S. A. (2013). Robust cultural pluralism: An interview with Professor Richard A. Shweder. *Europe's Journal of Psychology*, 9(4), 671–686. doi:10.5964/ejop.v9i4.718
- Skitka, L. J. (2014). The psychological foundations of moral conviction. In J. Wright & H. Sarkissian (Eds.), *Advances in Moral Psychology* (pp. 148–166). New York: Bloomsbury Academic Press.
- Skitka, L. J., & Mullen, E. (2002). The dark side of moral conviction. *Analyses of Social Issues and Public Policy*, 2(1), 35–41.
- Smith, I. H., Aquino, K., Koleva, S. et al., (2014). The moral ties that bind ... even to out-groups: The interactive effect of moral identity and the binding moral foundations. *Psychological Science*, 25(8), 1554–1562. doi:10.1177/0956797614534450
- Stewart, C. (1991). The Desert Fathers on radical self-honesty. *Vox Benedictina: A Journal of Translations from Monastic Sources*, 8(1), 7–54.

- Stojanov, K. (2012). The concept of Bildung and its moral implications. Paper presented at the Philosophy of Education Society of Great Britain, New College, Oxford. <http://www.philosophy-of-education.org/uploads/papers2012/Stojanov.pdf>
- Tillich, P. (1952). *The Courage to Be*. New Haven: Yale University Press.
- Vaillant, G. E. (2015). *Triumphs of Experience: The Men of the Harvard Grant Study*. Cambridge, MA: Belknap Press.
- van Gelder, J. L., Hershfield, H. E., & Nordgren, L. F. (2013). Vividness of the future self predicts delinquency. *Psychological Science*, 24(6), 974–980. doi:10.1177/0956797612465197
- vanDellen, M. R., Shah, J. Y., Leander, N. P. et al., (2015). In good company: managing interpersonal resources that support self-regulation. *Personality and Social Psychology Bulletin*, 41(6), 869–882. doi:10.1177/0146167215580778
- Vignoles, V. L. (2019). Identity: Personal AND social. In K. Deaux & M. Snyder, *Oxford Handbook of Personality and Social Psychology* (pp. 289–315). Oxford: Oxford University Press.
- von Bingen, H. (1141/1982). *Ordo Virtutum*. Deutsche Harmonia Mundi.
- von Bingen, H. (1141/1990). *Scivias* (C. Hart & J. Bishop, Trans.). Mahwah, NJ: Paulist Press.
- Walker, L. J. (2004). Progress and prospects in the psychology of moral development. *Merrill-Palmer Quarterly*, 50(4), 546–557.
- Walker, L. J., & Frimer, J. A. (2007). Moral personality of brave and caring exemplars. *Journal of Personality and Social Psychology*, 93(5), 845.
- Walker, L. J., & Frimer, J. A. (2015). Developmental trajectories of agency and communion in moral motivation. *Merrill-Palmer Quarterly*, 61(3), 412–439.
- Washburn, A. N., & Skitka, L. J. (2017). Science denial across the political divide. *Social Psychological and Personality Science*, 9(8), 972–980. doi:10.1177/1948550617731500
- Weidman, A. C., Cheng, J. T., & Tracy, J. L. (2018). The psychological structure of humility. *Journal of Personality and Social Psychology*, 114(1), 153–178. doi:10.1037/pspp0000112
- Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2013). The integrative process: New research and future directions. *Current Directions in Psychological Science*, 22(1), 69–74. doi:10.1177/0963721412468001
- Williams, B. (1981). *Moral Luck*. New York: Cambridge University Press.
- Wright, J. C. (2021). Morality as a regulator of divergence: Protecting against deviance while promoting diversity. *Social Cognition*, 39(1), 81–98.
- Wright, Z. V. (2015). *Living Knowledge in West African Islam: The Sufi Community of Ibrahim Niasse*. Leiden: Brill.
- Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, 21(3), 253–277. doi:10.1177/1088868316652279

Coda: Taking Moral Action

one day I went past the memorial plaque which had been put up for Sophie Scholl in *Franz Josef Strasse* [in Munich], and I saw that she was born the same year as me and she was executed the same year I started working for Hitler. And at that moment I actually sensed that it was no excuse to be young, and that it would have been possible to find things out.

(Traudl Junge in *Blind Spot: Hitler's Secretary*, 2003)

This intimate story of recognition of moral failure comes from Hitler's personal secretary, Traudl Junge, who worked in this role from 1942 (the year that resistance fighter Sophie Scholl was executed) until his suicide in the Berlin bunker in 1945. She wanted to be a ballet dancer but was eager to apply for the secretarial position in the Chancellery so that she might be at "the source of all information." After her capture, she was interrogated often by the Russians and Americans but then allowed to live freely in Bavaria until her death in 2002. She wrote about her experiences (Junge, 2005), claiming to have been ignorant of the great crimes of the regime until she was told about them after the war. Her comments in the earlier quote come from a documentary film based on extensive interviews with her that was released in Berlin the day before her death (Heller & Schmiderer, 2003). Its title comes from her recognition that rather than being on the inside, she was "in a blind spot." Shortly before her death, Traudl Junge said to the producer of the documentary, "Now that I've let go of my story, I can let go of my life" (Hooper, 2002, February 14).

This story crystalizes many of the themes of this book. It contains within it the struggle and moral heroism of Sophie Scholl and her colleagues to call out the true nature of the regime. Traudl Junge's later writing and speaking about her experience is in part an attempt to construct a narrative to excuse herself but also a recognition of moral failure and a call for us to "find things out" about our own complicity in injustice. Her narrative and confession place her firmly in the Gray Zone described by Holocaust survivor Primo Levi, the complex space between the simplified poles of good and evil (Levi, 2017). This linking of Sophie Scholl and Traudl Junge across the decades shows the intricate ways that individual choice, personal integrity and blindness,

openness to personal and cultural critique, and the changing moral ecosystems over time all interact to shape moral action.

Across the nine chapters and a wide variety of empirical and theoretical work, in this book we have been seeking an understanding of the **contexts** (*evolution* (Chapter 1), *neuroscience* (Chapter 2), and *moral ecology* (Chapter 3)), **influences** (*personality* (Chapter 4), *moral identity and the self* (Chapter 5), and *skills and knowledge* (Chapter 6)), and **processes** (*moral reason* (Chapter 7), *moral emotion* (Chapter 8), and *moral formation* (Chapter 9)) leading to moral action. In each chapter we have discussed research and literature we regard helpful to navigate the complexities inherent in a more inclusive landscape of moral psychology, grounded in how each person navigates their way among the multiple commitments and constraints of the world into which we have been thrown. In this short Coda, we first revisit the initial themes from the Introduction to see how they have been fulfilled. We then present four important concepts that have emerged from our journey through the voluminous literatures reviewed in each chapter.

C.1 Revisiting Our Initial Themes

In the Introduction, we highlighted four themes that would weave the text together: (1) a broad search for *variety in moral action*, influences, and processes (2) a concentration on the *idiographic experience*, choice, and development of the individual, (3) a focus on the ultimate, *normative status of moral claims*, and (4) a hope toward *educational implications*.

C.1.1 Variety in Moral Action

We have surely discovered variety along the way. The text has been one long endeavor to find some structure in this rich complexity of variation. The wide range of contexts, influences, and processes in the chapters themselves helps to bring some structure to the variety. And in cases, examples, and traditions of moral formation we have found variation in the content and definition of moral action ranging from the more conventional dedication to social justice or compassionate service, to passionate dedication to religious or ultimate truth, to self-transcendent involvement in a multitude of specific commitments and projects. Any claim to general processes or theory in moral psychology will need to grapple with this complexity and variation.

C.1.2 Idiographic Experience

In retrospect, most chapters (excluding Chapters 1 and 2) have been relentlessly focused on the experience of the individual in their role as agent. In more technical terms, they have been idiographic. This arises in part from the focus on variety, but it is also inspired by a continental philosophical take on the existential problems of everyday living and the need for the individual to hold this together in tension with the pull of the ultimate. Chapters 4, 5, and 6 present this turn to the idiographic in some detail. Even Chapter 3 frames the individual as a source of critique within moral ecologies.

C.1.3 Ultimate Status of Moral Claims

We have tried, in addition, to foreground the special characteristic of the ultimate in our view of morality, e.g. in our discussion of existential appropriation and self-transcendence in Chapter 9 or cultural difference in Chapter 3. This characteristic is a central aspect of the moral domain and connected closely to the normative “should.” It is that aspect of the normative that helps to put in place the multiple values any individual holds. Our overview and conclusions are still compatible with moderately complex versions of moral relativism, but they are more comfortable with pluralism, the idea that there are broadly shared multiple values, based in the human condition, that shape our moral commitments. These are sometimes in tension with each other, but even this tension can be resolved by creative integration of one as a means to the other. This process of integration is captured in the existential turn we have woven throughout the book, in which we present the individual as an intentional being who longs for ultimate meaning and is able to self-transcend (Frankl, 1959, 1997), pulled by “ultimate verities” (Maslow, 1969), or that which “concerns us ultimately” (Emmons, 1999; Hoffman & Ellis, 2019; Tillich, 1951).¹

C.1.4 Educational Implications

The applied aspect of our project has given rise to discussions of the development of skills and of moral formation in the individual, but also of moral ecology as a design project and an aspect of friendship (see Chapter 3). In this way we have transformed our educational hopes from passive learning to the active involvement of the individual in constructing a moral life and a moral ecology that supports it by facilitating the process of moral formation.

C.2 Toward a Theoretical Framework for Moral Psychology

In addition, along our journey through the chapters we have developed some concepts that offer a framework for understanding the wide variety of patterns and complexities in moral action. The chapter structure itself provides an expansion of what we need to understand to grasp the world of moral action. In each chapter we have occasionally reformulated the categories for that literature, and this re-sorting has often resulted in new insights. See, for example, Table 5.1 in Chapter 5. The four ideas we present here are crucial to the framework we want to construct for any adequate psychological understanding of moral action. They are based in the most recent research, but most of them have millennia-long traditions behind them and can be found woven into the fabric of the three traditional accounts of moral formation we offer in Chapter 9 (*Bildung*, existential appropriation, and monastic/spiritual formation). We hope this short list allows the reader to see the outlines of a pattern in the vast complexity and uncertainty of the human moral condition.

¹See Introduction, Section I.6.1.1.

C.2.1 A Central Process: Critique

Simply put, there are no independently verifiable ideal moral foundations that survive the rough and tumble, complex struggle of human existence in all its variety. Thus, the central question here is not: *What are the sure foundations of right action?* We are left instead with a question like: *How do I navigate this particular life?* We are left to choose among and to commit to the various projects, ideals, and values available to us in our human condition, constrained and supported as we are by particular moral ecologies and by the capacities we inherit and develop. This is the unavoidable existential task each individual is given.

But we do have support and resources for navigating our way and critiquing not only our own path, but the paths of others and of the moral ecologies in which we are embedded. We can choose to engage in personal, social, and cultural critique using the resources we have and the ones available in our moral ecology, narratives that speak to the “ultimate verities,” such as “truth, goodness, beauty, perfection, excellence, simplicity, elegance” (Maslow, 1969, p. 4). And we can take moral action at multiple levels as critique of our culture. Critique is only one form of moral action, but all moral critique is moral action, whether quietly internal to the self or boldly public and social. As a scientific enterprise, moral psychology can help to identify and describe the resources, capacities, constraints, and processes of this critique on multiple levels.

C.2.2 A Framework for Moral Evaluation and Commitment (SERE)

To engage in critique, one needs some criteria to evaluate one’s own actions, others’ actions, and the influences of one’s moral ecology. The primary question here is: *How do I reflect and decide on moral action?* Any reasonable model of moral evaluation will need to include the three aspects of social–emotional–reflective equilibrium. Social refers to the social embeddedness of the evaluation and its dependence on the resources and constraints of one’s particular moral ecology. Emotional–reflective is based in the deep interconnection of emotion and reason that is a unique aspect of the human condition. The word reflective should embrace both explicit and implicit evaluation since much moral action is likely not consciously reflective but based in long-standing commitments, deeply practiced skills, and cultivated habits. The three aspects together provide some balance to motivate and guide our evaluations both immediate and long term. Again, there is no unquestioned foundation here, but simply a recognition that the best one can do is to seek an appropriate equilibrium among these three aspects at different times or domains in one’s life.

We do not propose a model for how this happens. We are simply specifying the kinds of things we consider when evaluating our actions and our commitments. The phrase reflective equilibrium is taken from philosophy and suggests a balance among considerations. But clearly if one is evaluating one’s ultimate concerns, equally weighting all the concerns is out of the question. How does one find a weighted balance in the face of the ultimate? Perhaps it is achieving the constant bridging of the tension instead (see the suggestions on this from the tradition of existential appropriation in Chapter 9). Recognizing disfluency in action, when the tensions become too great (see Chapter 6) helps one to break out of automatic processes and begin conscious, controlled re-evaluation of a host of things, including one’s ultimate

commitment (see Chapter 5 for the complexity involved here). One's moral ecology can support the recognition of disfluency or suppress and defend against it. We describe these influences in Chapters 3 (see Section 3.3.2) and 6 (see the distinction between wicked and kind environments there, Section 6.3.1). This approach does not solve the problems of the human condition or the difficulties of moral choice. Still, understanding the complexity of self-evaluation and commitment in the search for integrity can provide resources for building this firmly anchored but constantly shifting equilibrium.

One crucial point on the stakes associated with the choices here: commitment to a self-transcendent goal has a dark side (Hart, 2005; Skitka, 2014; Skitka & Mullen, 2002). Ultimate commitments to some values, causes, or projects can produce terrorism and other forms of evil (Koltko-Rivera, 2006). The psychological processes involved in this look a lot like self-transcendence. There is a great deal of work needed to distinguish between the goals. It may be that the psychological processes are very similar if not identical. Recognizing this dark side allows us to frame the crucial and unavoidable work of distinguishing what the good is and whether how we become good is the same process as how we become evil (see the discussion of "double reflection" in Chapter 9, Section 9.1.2).

None of these aspects or tensions within SERE are new ideas, and their combination is simply one that has been left fallow. One can find its roots in major philosophical and religious traditions across cultures. It is, in fact, part of our evolutionary inheritance of extensive interconnection between emotion and reason dedicated to guiding action in a social creature complex enough to long for meaning and morality. It is our peculiar existential task to form it in ourselves.

C.2.3 Three Aspects to Moral Formation

There are three aspects to how we get better at taking moral action. One is *evaluative* and primarily asks the question: *How am I doing?* It thus involves recognizing relevant situations and internal states and processes, evaluating them in light of goals, values, or standards, and making and implementing plans for change in a desired direction. It is likely within evaluative moral formation that the SERE model of moral evaluation is primarily active. For, as emphasized in Section C.2.1, personal moral reflection *is* moral action.

The second aspect is *normative*, and asks primarily: *Where am I going?* This is a question of what goals, or what kinds of goals, I should commit myself to. One can see this formation in the transformation of goals across a lifetime or through critical periods. One can also find it happening in the value tradeoffs that one must constantly negotiate in the various domains of life. Here the SERE model is again relevant.

The third aspect is *educative* and asks: *Where can I find support and guidance?* This is a personal question of how one can find moral resources and a collective question of how we can construct environments that support moral formation.

C.2.4 A Dimensional Landscape for Moral Formation

We began this book with a Humboldtian emphasis on variation. Humboldt's science was driven in part by geography, finding the places where undiscovered variation, and

patterns in variation, occurred. Our search for variation in moral action initially produced the list of chapters detailing the context for, influences on, and processes of moral action. We then added some cases to stretch our conceptions and reflected on three traditions of moral formation in a way that allowed for a dialog between theory and application.

In Chapters 3 and 7, we found a variety of models of the abstract value space within which variation in moral action might be discovered. Then, in Chapter 9, we found that at the concrete level of individual moral action variation exploded into a multitude of individual commitments and projects (e.g. caring for these particular children or that social justice issue, or this personal religious call). In the end, we have settled on a value landscape that includes three dimensions as a way to find some structure in the huge multiplicity of value commitments: self-transcendence, openness, and inclusion, or TOI.

C.2.4.1 Transcendence, or What is It About? The first dimension, transcendence, is taken from work by Schwartz (2012, 2016) and has commitment to personal agency at one end and commitment to a self-transcendent goal at the other.² The central question for this dimension is: *What is my morality about?* This is the one dimension for which one can make a clear developmental argument: that we necessarily begin in childhood at the personal agency end and become more competent members of society by incorporating some balance with self-transcendent goals, particularly care for others (Walker & Frimer, 2015). But even here there is interesting variation, with moral exemplars folding the dimension in on itself, fusing the agency and transcendence poles by endorsing agency in service of the transcendent goal. And one also finds a variety of possible transcendent goals: care, social justice, religion, art, etc. (Koltko-Rivera, 2006; Maslow, 1969).

C.2.4.2 Openness, or How Am I Doing It? Perpendicular to the transcendence dimension in all the empirically derived models is a conservation to openness dimension (see Zapko-Willmes et al., 2021 for a review). The central question for this dimension is: *How do I go about doing my commitment?* The individual goals in the transcendence dimension can each be approached in a variety of ways, some more conservative/traditional and others more open to social change. One can care for the environment by conserving particular ecosystems or by changing social and economic structures. One can care for the poor through direct service or through social justice campaigns. Allowing for this variation in our landscape allows us to expand the variety of moral exemplars we might look for.

²Labels on dimensions like these are always fluid. In their adaptation of Schwartz (2016), Walker and Frimer (2015) replace “achievement” with “agency” and “self-transcendence” with “communion” as their labels for the endpoints of this dimension, in part to connect it to other literatures. We prefer self-transcendence to communion in order to emphasize the variety of non-social goods to which one might be morally devoted (religion, art, environment, etc.) and thus to take seriously this variety of the pull of the ultimate.

C.2.4.3 Inclusion, or *Who is It for?* The third dimension is about who is included or excluded from our circle of moral concern. The question here is: *Who is my moral commitment for?* One need not think of this in terms of ethnocentrism *vs.* inclusion, though this is one obvious way to use this dimension (see Chapter 9, Section 9.3.1.2). One can also think of the focus of concern being narrow (these particular children or guests) or broadly inclusive. And the scope of the moral circle might change over time in response to external threats (Gelfand, 2012, 2021; Wright, 2021). Again, this dimension might also allow us to search for a broader sample of moral exemplars.

We have lightheartedly named this three-dimensional space TOI (transcendence, openness, inclusion, pronounced “toy”) to suggest its tentative nature as a tool to discover variation in moral action. Thus, it is not a comprehensive model of moral action or development but rather a guide to exploration of variation in moral action.

C.3 An Adaptable Framework

These suggestions toward any theory of moral action are tentative, exploratory, and adaptable. There is ample room for multiple competing and cooperating theories in the structural constraints and variation we outline. But any adequate theory will need to account for both the structure and the variation or at least justify its own boundaries for the field. Also, any adequate theory will need to propose how development, (trans)formation, and change happen personally and within cultures, and thus necessarily entangle itself in normative questions. Finally, adequate theory will need to grapple with the complexity and depth of the real world at least as represented in the incomplete range of cases and traditions we have offered here. We present this framework in the hope that it will contribute to thoughtful, interdisciplinary attempts to grapple with the science and story of moral action.

References

- Emmons, R. A. (1999). *The Psychology of Ultimate Concerns: Motivation and Spirituality in Personality*. New York: Guilford Press.
- Frankl. (1959). *Man's Search for Meaning*. New York: Washington Square Press.
- Frankl. (1997). *Man's Search for Ultimate Meaning*. New York: Insight Books.
- Gelfand, M. (2012). Culture's constraints: International differences in the strength of social norms. *Current Directions in Psychological Science*, 21(6), 420–424. doi:10.1177/0963721412460048
- Gelfand, M. (2021). Cultural evolutionary mismatches in response to collective threat. *Current Directions in Psychological Science*, 30(5), 401–409. doi:10.1177/09637214211025032
- Hart, D. (2005). Adding identity to the moral domain. *Human Development*, 48, 257–261.
- Heller, A., & Schmiderer, O. (Writers). (2003). *Blind Spot: Hitler's Secretary*. D. Krausz & K. Stocker (Producer): Sony Pictures Classics.
- Hoffman, L., & Ellis, M. S. (2019). Ultimate concern. In D. Leeming (Ed.), *Encyclopedia of Psychology and Religion* (pp. 1–4). New York: Springer.
- Hooper, J. (2002, February 14). Traudl Junge. *The Guardian*.

- Junge, T. (2005). *Until the Final Hour: Hitler's Last Secretary* (A. Bell, Trans. M. Müller Ed.). New York: Arcade.
- Koltko-Rivera, M. E. (2006). Rediscovering the later version of Maslow's hierarchy of needs: Self-transcendence and opportunities for theory, research, and unification. *Review of General Psychology*, 10(4), 302–317. doi:10.1037/1089-2680.10.4.302
- Levi, P. (2017). *The Drowned and the Saved*. New York: Simon & Schuster.
- Maslow, A. H. (1969). The farther reaches of human nature. *The Journal of Transpersonal Psychology*, 1(1), 1–9.
- Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2(1). doi:10.9707/2307-0919.1116
- Schwartz, S. H. (2016). Basic individual values: Sources and consequences. In T. Brosch & D. Sander (Eds.), *Handbook of Value* (pp. 63–84). Oxford: Oxford University Press.
- Skitka, L. J. (2014). The psychological foundations of moral conviction. In J. Wright & H. Sarkissian (Eds.), *Advances in Moral Psychology* (pp. 148–166). New York: Bloomsbury Academic Press.
- Skitka, L. J., & Mullen, E. (2002). The dark side of moral conviction. *Analyses of Social Issues and Public Policy*, 2(1), 35–41.
- Tillich, P. (1951). *Systematic Theology, Vol. 1, Reason and Revelation. Being and God*. Chicago, IL: University of Chicago Press.
- Walker, L. J., & Frimer, J. A. (2015). Developmental trajectories of agency and communion in moral motivation. *Merrill-Palmer Quarterly*, 61(3), 412–439.
- Wright, J. C. (2021). Morality as a regulator of divergence: Protecting against deviance while promoting diversity. *Social Cognition*, 39(1), 81–98.
- Zapko-Willmes, A., Schwartz, S. H., Richter, J. et al., (2021). Basic value orientations and moral foundations: convergent or discriminate constructs? *Journal of Research in Personality*, 92(June), 104009. doi:10.1016/j.jrp.2021.104099

Index

Please note that page references to Figures will be followed by the letter ‘f’, to Tables by the letter ‘t’. References to Notes will contain the letter ‘n’ followed by the Note number

- abstract knowledge x, 267
- ACC (anterior cingulate cortex) 38
- accountability 128, 130
- action
 - active suppression of 152
 - defining viii
 - guiding of 152–153
 - moral self-contempt influencing 122–130
 - see also* moral action
- action tendency 13, 220
- adaptation
 - adaptive behavior 10–12
 - adaptive personality 90–92
 - adaptive sociality 12–13
 - characteristic adaptations 35, 98
 - evolutionary 5, 10, 12
- adolescents 40, 270, 282
- see also* children
- Adolphs, R. xxi–xxii, 15n.12
- Advaita Vedanta Hinduism 254
- advantageous comparison 200
- agency 18, 105, 124, 218
 - agentic values 270
 - categorizing, 266n.32
 - displacing with communion 270, 282
 - and moral ecology 64, 68
 - and moral formation 257, 268, 277
 - and neuroscience 34, 39
 - personal 296
 - self-agency 132
 - in service of others 46
 - supernatural 232
- aggression viii, 10, 22, 75, 95, 200, 217, 223, 229
- agreeableness–disagreeableness personality
 - dimension 36, 100
- Ainsworth, M. 121
- Allport, G. ix
- Alter, A. L. 152–153
- altruism
 - vs. egoism xxi
 - and empathy 22, 36
 - gene–culture coevolution 7
 - and the pISTS 36
 - vs. selfishness 9–10
- altruistic punishment 3, 7, 8, 22, 222, 224
- Ambady, N. 32
- amygdala 30, 38, 43
- Anand, V. 66, 67
- anger 31, 216n.2, 219n.6, 220, 223–224
 - appraisal theories 216, 217
 - cultural theories 218
 - and disgust 44, 73, 213
 - against injustice 153, 222
 - righteous 117, 221
 - vicarious 20, 37n.16
 - see also* aggression; disgust; emotions; fear; moral emotions
- animals *see* non-humans/primates
- anterior cingulate cortex *see* ACC (anterior cingulate cortex)
- anterior insula 43
- anticipatory fear, of punishment 3, 13
- anti-Semitism xiii, 107
- Apostle Paul 115, 116
- Appiah, A. 250
- appraisal theories of emotion 216, 217–218, 219
- appropriation
 - cultural 251n.9
 - goal 248, 252
 - subjective 253n.17
 - of ultimate good 271
 - value–goal–project process 271
 - see also* existential appropriation
- Aquinas, T. 157n.25, 219n.7
- Argentina, dictatorship of 1976–1983 xii

Taking Moral Action, First Edition. Chuck Huff and Almut Furchert.
© 2023 John Wiley & Sons Ltd. Published 2023 by John Wiley & Sons Ltd.

300 Index

- Aristotle 41, 102, 107n.22, 136, 156n.24, 163, 168, 223n.14
Nicomachean Ethics 163, 246
Politics 3
- Armstrong, K. 226
- Arnold, M. 157n.25, 219n.7
- Ashforth, B. 66, 67
- Ashley, K. D. 162
- Atkins, R. 91
- attraction–selection–attrition cycle 68–69
- Augusto of Hippo 189
- authentic pride 228, 229
- automated response dispositions 163
- Avolio, B. J. 155
- awe 231–232
- Bandura, A. 67, 200
- Bargh, J. A. 161, 190n.10
- Barnard, L. K. 62, 72, 92, 99–100, 119, 122, 263n.28, 264–265
- Barrett, L. F. xixn17, 215, 216n.1, 219
- Basso, J. C. 34, 39
- Bates, T. C. 6
- BB synchrony *see* brain-to-brain synchrony (BB synchrony)
- Beall, J. 126n.21
- Bebeau, M. J. 122, 188, 264
- Beer, J. S. 37, 41
- behavior
 acceptable standards of xv
 adaptive 10–12
 of ancestors 15
 angry 223
 anti-social 22, 31
 behavioral signatures 91, 95, 96
 changing 154, 164
 and character, xxiii 90
 compassionate 9
 complex 31
 consistency of 91, 125, 128, 165, 246
 control over 147, 163
 cooperative 6, 7, 9, 13, 20, 45
 criminal 99, 100
 decision-making 167
 eating 164
 ethical 68, 69, 72, 78
 evolutionary theories 8, 9, 13–15, 19
 expression of 148
 external 123, 124
 future 120, 257
 group 93
 imitating 32
 immoral xvii, 78
 improper 228
 labeling 103
 legal/illegal 43
 mindfulness changing 154
 mirroring 14
 moral viii, ix, xvii, xxiv, 8, 29, 76, 122, 124, 127, 131, 133, 155, 162, 163, 270
 moral formation 256
 motivating 107, 221
 and narrative 257
 non-verbal 45
 nurturing 90
 organizational misbehavior 66
 patterns of 96
 and personality xxiv, 91, 94–96, 148, 269
 predicting 120
 pro-social ix, 13, 21, 22, 90, 104, 109, 221, 232
 and restorative justice 107
 search for consistency in moral behavior viii
 and self-concept 124
 and shame 230
 skilled, xxxi 146
 social 11, 29
 socially responsible 121, 264
 strategies 9
- behavioral integrity 124–125, 127, 133
- Bentham, J. 95n.9
- Bharucha, J. 32
- bias 117n.10, 196, 204
 moral emotion as 42–43
 vs. reason xxiv–xxv
- Biden, J. xiii
- Big Five structure of personality 36, 93, 94, 99, 106
 alternatives to the Big Five 101
 moral correlates of the Big Five 99–101
see also personality
- Bildung* (education as personal formation) 248, 249–250, 255, 258, 271, 275, 276
see also monasticism/monastic and spiritual formation; religion
- blame judgments xvii
- Blasi, A. 116, 117, 122, 123
- Bodolica, V. 229
- body awareness 18
- Bonhoffer, D. 127
- Borenstein, N. xi, xiv–xv, 77, 127, 166, 204–205
- Borg, I. 270n.40
- Bowlby, J. 121
- Bowman N. A., 62
- Boyd, R. 7
- Brackett, M. A. 233–234
- brain
 cultural, group and individual influence on processes of 33–34
 and cultural processes 34–35
 imaging of 31 33, 36
 increase in size 3
 lack of moral “center” in 46, 47, 49
 mirror neurons 32
 moral 39, 47
 “organ of culture” 32
 social 3–4
 specific pathways or structures in, xx
see also neuroscience
- brain-to-brain synchrony (BB synchrony) 34, 45–46
- Braithwaite, J. 67
- Brief, A. P. 58

- Browning, C. R. 90n.2
 Buber, M. 226n.18
 Buddhism 226n.20, 253, 254, 258
 building blocks, evolutionary 12–15, 21
 Buttram, R. T. 58
- CAM *see* central affiliation mechanism (CAM)
 Campitelli, G. 160–161
 CAPS *see* cognitive-affective processing system (CAPS)
 care, concept of 187–188
 Carnes, N. C. 265
 Carver, C. S. 229
 Casey, M. 254n.20
 Cawley III, M. J. 101, 102
 central affiliation mechanism (CAM) 6–7
 Cervone, D. 94, 97, 98, 118, 127n.23
 Cesario, J. 198
 character 89, 90
 Cheney, E. xiv, xvi, 253, 271–272
 Cheng, J. T. 232n.33, 254
 Chicago school of sociology 58
 children 19, 45, 73, 95, 100, 162, 168, 297
 challenges/disabilities, with 20, 21, 90, 104, 132, 247, 279
 empathy for xii, 21, 90
 fostering of xii, xx, 20, 132, 247, 264, 278
 and honesty xxiii–xxiv, 116, 269
 missing 150
 moral emotions 229, 233
 moral formation 273, 296
 other-oriented 188
 under- or over-controlled 91n.3
 resilience in 91n.3
 self-conscious 18
 tracking 95
 see also adolescents
 Chinese culture 33, 34
 Christianity 61, 92, 134, 249, 254, 258
 monasticism 275
 Rule of Benedict 232n.33, 273
 see also *Bildung* (education as personal formation); religion
 “chronics” 158
 civil rights movement xiii, 266n.32, 275
 class differences 62
 Clayton, N. S. 15n.12
 cognition
 cognitive-affective processing system 95–97
 dual process theories 189–190, 191*t*
 and emotion
 planning for cognitive-emotional effectiveness 41–42
 questioning the cognitive/emotion distinction 44
 implicit cognition and two-process models 189–198
 naturalistic moral 198–199
 see also cognitive control; moral cognition
 cognitive appraisal 13
 cognitive chronicity 155
 cognitive conceptualization 220
 cognitive control 17, 40
 cognitive entrenchment 164
 cognitive network metaphor 118
 cognitive-affective processing system (CAPS) 95–97
 Cohen, A. B. 61
 Cohen, J. D. 41, 42, 45, 215
 Colby, A. 62, 66, 72, 74, 92, 105, 126, 163, 183, 258, 259, 266, 273–274, 277n.49
 collectivist cultures 37n.16
 community
 close relationships in 69–71
 community concern 9, 13, 21, 57
 community service xii, 122, 159, 264, 266
 morality, community-based 70
 compassion xx, xxv, 4–6, 9, 14, 20–22, 39, 44, 47n.28, 75, 196, 216n.2 222, 254, 259, 262
 and agreeableness 100
 class differences 62
 deservingness 227
 and embarrassment 228n.25
 emotional consistency 126
 emotional regulation 153
 and human rights 264
 individual differences 227n.22
 as a “master social emotion” 227
 meditation-based training 154, 233, 235, 237
 monastic traditions 271
 moral emotions 226–227
 moral exemplars 263
 self-compassion 226n.20
 social-emotional reflective equilibrium (SERE) 236–237
 see also empathy
 conflicts of interest 65
 conformity 183, 184
 Confucius 168
 conscience 13–14
 conscientiousness–unconscientiousness personality dimension 36, 94, 100
 consciousness
 Kohlbergian paradigm of ix
 self-consciousness 18
 social 18
 consistencies
 of behavior 91, 125, 128, 165, 246
 consistent moral action, drive to unity supporting 134
 cross-situational xxiv, 116
 emotional 126
 evaluating xxx
 instrumental 126–127
 logical entailment 125, 126
 moral 96, 108, 125
 multiplicities of consistency 135
 narrative 125n.20, 127
 search for consistency in moral behavior viii
 social 126
 types of 125–128, 129*t*
 and variable self 120
 constructionist models of emotion 216, 219–220

302 Index

- contempt 224–225
 - see also* moral contempt
- cooperative behavior 6, 7, 9, 13, 20, 45
- Corbin, A. 218
- Cosmides, L. 193
- culture
 - brain and cultural processes 34–35
 - Chinese 33, 34
 - contexts viii
 - cultural anthropology 34
 - cultural neuroscience 34
 - cultural pluralism 188
 - cultural theories of emotion 218–219
 - culturally pervasive ideas 59
 - vs. evolution 7–8
 - honor/shame system in Middle Eastern countries 61, 187
 - human distinctiveness 17
 - humans as active co-creators of influence in 77
 - and moral action xvii
 - and moral diversity 59–65
 - patterns across cultures 63–65
 - Western 33–34, 99, 230
- Curry, O. S. 63
- cynicism 105
- Dahl, A. 198, 205, 275n.46
- Damon, W. 62, 66, 72, 74, 92, 105, 126, 163, 258, 259, 266, 273–274, 277n.49
- Daniels, N. 202n.21
- Darley, J. M. ix, 148n.10
- “Darley’s Law” 68n.9
- Darwin, C. xix, 6, 13, 219
 - The Origin of Species* 4
- Dautheribes, T. M. 63
- Davenport, J. 147n.6
- Davis, M. 65
- Dawkins, R. 9
- De Raad, B. 101–102
- De Ridder, D. T. D. 151
- De Vries, R. 105
- De Waal, F. B. M. 9, 13, 15, 150n.14
- deception/self-deception 18, 115, 268n.37
- decision theory 180
- decision-making xvi, 65, 251
 - cognitive/emotion distinction, questioning 44
 - and decision theory 180
 - evaluation in 155
 - moral xxxi, 49, 146, 167
 - and moral judgements 39
 - naturalistic 198, 199
 - principled 167
 - processes 71
 - and reason 180
 - self-interested economic 40
 - situational influence 79, 109
 - supporting 42
 - two-process models 192
- deference to authority 13, 194, 197
- dehumanization 201
- design issues
 - ethics as 275
 - moral action as 167
 - moral ecology as xxx, 57
 - morality as 76
 - organizational ethics as 71–72
- determinism 31
 - vs. free will xxi, 150n.14
- development
 - biological 185
 - brain *see* brain; neuroscience
 - conceptual 146n.2
 - emotional 126, 154
 - Greek ideal 251n.10
 - individual xxviii, 206, 261
 - moral *see* moral development
 - of non-egocentric self 253, 254
 - of personality 109
 - self-development 159
 - and self-regulation 14
 - and self-transcendence 19
 - sensitive periods for 161
 - vs. stability xxiv
 - and theology 250
 - of virtues xxiii, 254
- developmental psychology vs. moral psychology ix, 181
- Dewey, J. 183–184
- DeYoung, C. G. 35, 36
- Dickinson, E. xvi
- dilemma-based assessment model 186
- Dilthey, W. xxviii, 28
- directed arousal 13, 220
- disgust 218
 - and anger 44, 73, 223, 224
 - anger–disgust–contempt triad 227
 - anticipatory 41
 - and contempt 225
 - and guilt 225
 - malleability 223
 - moral 223
 - sexual 223
 - see also* anger
- dispositional moral traits xxx, 90
 - agreeableness–disagreeableness 36, 100
 - alternatives to the Big Five 101
 - conclusions about 103
 - conscientiousness–unconscientiousness 36, 100
 - extraversion–introversion 36, 99–100
 - mapping virtues as traits 101–103
 - moral correlates of the Big Five 99–101
 - neuroticism–emotional stability 36, 100–101
 - openness–closedness 101
- domain limitations 66, 67, 75, 123
- domain specificity/domain-specific skills 146, 156, 163, 166, 168
- Dovidio, J. F. 99
- Downs, J. S. 180
- Dreyfus, H. L. 159, 160, 161, 165
- Dreyfus, S. E. 159, 160, 161, 165
- Dukerich, J. M. 58
- Durkheim, É. 69
- Durr, V. 274, 275

- ecological rationality 193, 202
 Eddington, K. M. 45
 EEG monitors 45
 egocentricity bias 46
 ego-depletion hypothesis 151
 egoism xxi, xxxiii
 vs. altruism xxi
 moral 130
 Eisthen, H. L. 198
 elevation and awe 73, 231–232
 Ellemers, N. ix–x, 38n.19 41, 46
 Elliot, M. 31n.5
 Ellsworth, P. C. 216n.1, 217n.3 218
 embarrassment 227–228
 embedded training systems xi–xii
 Emery, N. J. 15n.12
 emotional consistency 126
 emotional intelligence 234
 emotions
 appropriate or inappropriate evaluations
 of 234–235
 and behavioral integrity 125
 as bias 117n.10
 blending with reason 17, 41, 49
 categories, confusion within and between 236
 and cognition 94
 constructionist models 216, 219–220
 defining xxxi
 emotional regulation 128, 151, 153–155
 emotion-based conscious awareness 194
 and feelings 13
 and language 216, 219
 literature xix
 moral *see* moral emotions
 and morality xxxi, 220–221
 negative self-emotions 14
 outwardly focused 37n.15
 and rationality 17, 47n.28, 235, 236
 and relativism 221
 self-conscious 37, 38, 222, 227–230
 social 13, 20
 spontaneous 153n.20
 theories 215, 217n.3 219
 appraisal theories 216, 217–218, 219
 basic theories 216–217
 cultural theories 218–219
 overview of theories 215–221
 see also fear
 empathy 11, 13, 20, 45–46
 affective 104
 and altruism 22, 36
 avoidance 200
 for children xii, 21, 90
 cognitive 104
 dispositional 104
 egalitarian 264
 and emotional consistency 126
 in groups 121
 habituated 154
 in medicine, 42n.22
 and neuroscience 30, 33, 36
 other-oriented 104
 perceptual and cognitive components 36
 and personality 90, 92, 102, 104, 105
 processes 33, 36
 see also compassion
 empirical psychology xxii, xxvi, xxxii, 184
 Engberg, S. 266
 engineering xv
 ethics 63, 162
 firms 63, 65
 moral exemplars 156, 163
 National Society of Professional Engineers 265
 professions 156
 rules vs. goals 265
 safety considerations 265
 schools 63
 values 65
 Epley, N. 71, 205, 275
 Erikson, E. 105
 Eskreis-Winkler, L. 256n.21
 ethics vii, xxii, xxvn.24, 57, 63, 250
 Aristotle on 163
 assessment 186
 autonomous vs. heteronomous xxvn.25
 business 155
 codes of 68, 69, 265, 268
 and compliance 67, 73
 courses 145
 decision-making xvi
 as a design issue 275
 education 145, 235
 engineering 162
 ethical behavior 68, 69, 72, 78
 guidelines 145
 meta-ethics xxii, xxvii, 186
 and morality 186
 organizational 67, 68
 quandary 186
 reward and ethical behavior 69
 in science 105
 teaching of 235
 technology 155
 training 235
 universal principles 183
 virtue ethics 160n.29, 160n.29, 231n.30, 253
 euphemistic labeling 201
 Evans, J. S. B. T. 190
 evidence xxix, xxxiii, 4, 5, 14, 94, 154, 156, 198,
 220n.8 258, 268, 270
 attachment styles 232
 and automaticity 161
 for central affiliation mechanism (CAM) 6–7
 collecting 10
 of conceptual complexity 257
 cross-cultural 11, 14
 empirical 16, 94, 107
 genetic 11
 hunter-gatherer 11
 medical 11
 and moral ecology 61, 63, 75, 77
 and moral emotions 217, 219, 225n.17, 231, 233
 and moral identity 121, 123, 124, 126, 130,
 133, 134

- evidence (*cont'd*)
 moral reasoning 187, 193, 194, 195, 196
 networks of 14
 personality 90, 94, 95, 101, 103, 105, 107
 phylogenetic 11
 physiological 11, 219
 psychological 11, 204
 quality of 9
 scientific 269
- evil 57, 70
 acts 197
 becoming 295
 forms of 262–295
 goals 199
 good, and 20–253, 280, 291
 psychology of 75
- evolutionary psychology 10
- evolutionary theories viii, xxix, 3–28
 adaptation *see* adaptation
 and application 20–21
 basic processes, moral perspectives 4–10
 behavior 8, 9, 13–15, 19
 building blocks 12–15, 21
 cooperative behavior 6, 7, 9
 early evolutionary systems 198
 evolution vs. culture 7–8
 fit of evolved moral capacities with social and
 physical ecology 21
 gene–culture coevolution 7
 group selection 6–7, 9
 human distinctiveness 15–20
 kin (family) selection 5–6
 model of evolution of morality 21
 moral norms 14–15
 and morality 12–15
 natural law 19
 natural selection 4, 9
 open questions 21–22
 pro-social behavior strategies 13, 21
 selfishness and altruism 9–10
 sexual selection 5, 7
 social brain 3–4
 “survival of the fittest” 4
 see also Darwin, C.
- exapted (compassion for a wide group) 6n.5
- executive function 146–147
- existential appropriation xxxii, 248, 251–253,
 255, 259, 279, 293, 294
 what and how of 260f
 see also appropriation
- existential turn xxii, xxvi–xxvii, xxviii, 260, 293
- expertise
 advanced beginners 160
 automaticity in 161
 competence 160
 and cultural disagreement about moral action 169
 development of 160–161
 domain dependent 165
 domain specificity 163
 expert performance 157–163
 kind environments 160, 167
 mature moral functioning 158
 music analogy 157
 and novices 160, 162
 proficiency 160
 types of knowledge and skill 161–162
 and virtue 157
- explicit judgment 184–185
- extraversion–introversion 36, 99–100
- fear 31, 152, 229, 236, 258
 anticipatory, of punishment 3, 13
 in dominance hierarchies 13
 learned 12
 meaning of situation driving 31
 and morality 20
 social exclusion vs. physical harm 30, 45, 220
 see also anger; emotions; moral emotions
- Federal Sentencing Guidelines for Organizations,
 US 67
- Fischhoff, B. 180
- Fishbach, A. 256n.21
- Fiske, A. P. 196, 197
- Fiske, S. T. 225n.15
- Flanagan, O. 58, 92
- fMRI (functional magnetic resonance
 imaging) 31, 41, 220
- Fodor, J. 189
- food chain 58
- forced disappearance xii
- Frankfurt, H. G. 17, 44, 128
- Frankl, V. xxvi–xxvii, xxviii28
 Man's Search for Meaning xxvi
- free will xxxiii 150
 vs. determinism xxi, 150n.14
- Freedom Summer, participation in 267
- Freud, S. 94, 146, 189, 216n.2
- Frimer, J. A. ix, xvi, 62, 120n.16, 124, 127, 130,
 132, 134, 195, 258, 261, 266n.32, 266,
 270, 277, 280, 296n.2
- frontal cortex 36
- Furchert, A. xxvn25
- Gadamer, H.-G. x–xi, xxii, xxviii.28, 165n.36,
 267, 271n.41
- Garud, R. 68
- Gehman, J. 68
- Gemeinschaft* 61
- gender, and concept of care 187–188
- gene–culture coevolution 5, 7
- general skills 145–156
- generalized other 19
- generativity 105–106
- genes 4, 7
 selfish 9
- genetic encoding 7
- genetic selfishness 9
- genome 4, 5
- Gesellschaft* 60–61
- Gestalt tradition viii
- Gibbs, J. C. 61–62, 186
- Gilligan, C. 69n.10, 187–188
- goals
 individual variation in 62–63

- and organizational values 65
- pro-social, using 72
- Gobet, F. 160–161
- Goethe, J. van 249
- Goetz, J. L. 13
- Goldin, P. R. 43
- Graham, J. 263, 264n.29
- Grandmothers of the Plaza xii
- gratitude 232
- Gray, J. R. 35
- Gray, K. xxn.21, 193n.15
- Greene, J. D. 40, 41
- Gregg, G. S. 61, 187
- group selection 5, 6–7, 9
- groups
 - adaptive sociality 12–13
 - group organization viii
 - membership 121
 - moral action taken by 77–78
 - selection pressure between 6
 - vicarious punishment for in-group members 10
 - within group selection 6
 - see also* group selection
- guesser–knower paradigm 18
- guilt 184, 225–226
- Gustafson, S. B. 91
- habits 163–166
 - detaching from context 164n.34
 - domain limited 167
 - moderate strength 164n.33
 - performance context cues 164
 - regulation of 164–166
 - skills and knowledge 39
 - strong 164n.33
- Haeckel, E. H. P. A. 58
- Haidt, J. 8, 59, 60, 65, 73, 100, 193–195
 - on moral emotions 221, 222, 235
 - social intuitionist model 126
- Hampson, P. J. 157n.25
- Hannah, S., T. 155
- “Happy victimizer” phenomenon 233
- Hardy, S. A. 276n.48, 278
- Hargraves, I. 252n.16
- Harris, C. 265
- Hart, D. xxv, 69–70, 91, 122, 199–200, 264
- Hartnagel, F. xiv
- Hartshorne, H. xxiii, xxiv, 99, 116, 269
- Hassin, R. R. 161
- hate crime xiii, 107
- Healy, P. 67
- Heatherston, T. F. 37
- helpfulness 104
- Helton-Fauth, W. 58
- Herd, J. 249, 250, 278n.51
- hermeneutics/hermeneutical circle viii, xi, 252
- heteronomous ethics/morality xxvn.25, 183
- hierarchy of needs xvi
- Higgins, E. T. 15, 38
- Hildegard, Benedictine Abbess 253–254
- Hill, P. L. 92, 97–98, 103
- Hill, R. 236
- Hindu culture 216n.2
- hippocampus 30n.2
- Hirschman, A. O., xvii 72–73
 - Exit, Voice, and Loyalty* 72
- Hirsh, J. B. 100
- historical perspective, moral psychology viii–x
- Hitler, A. 127, 291
- Hofstede, G. 59, 64, 195n.18
- Hogg, M. A. 74
- Holocaust xv, 90n.2, 92, 121, 156, 227, 263, 291
- Homo economicus 9
- honesty xxiii–xxiv, 94, 116, 269
- hubristic pride 228, 229
- Huff, C. xxvn.25, 58, 62, 72, 92, 99–100, 119, 122, 263n.28 264–265
- Hulsey, T. L. 157n.25
- human distinctiveness 15–20
 - cognitive control and self-regulation 17
 - extensive culture 17
 - extensive integration of reason and emotion 17
 - manipulation of others’ perception 18
 - reality, sharing with others 19
 - religious experience 19
 - self-actualization 19
 - self-consciousness and self-awareness 18
 - self-transcendence 19–20
 - social consciousness 18
 - symbolizing and reasoning 16
 - time perspective 16
 - universalized normative evaluations 19
- human ecology 58
- human rights
 - in Argentina xii
 - basic 250
 - and compassion 264
 - language 134, 250, 274
 - universal xiii
 - in the US xiii
- Humboldt, W. von 249, 272, 295–296
- Humboldtian Science xix
- Hume, D. 179, 182, 192
- humors 35
- IAT (implicit association test) 43
- identity, moral *see* moral identity
- identity integrity 123–124, 128
- idiographic experience, xxvii, xxviii 149, 292
 - idiographic vs. nomothetic approaches 93–95, 98
- imitation 32
- implicit association test (IAT) 43
- incest avoidance/incest taboo 8, 10–12, 14, 194
- inclusion
 - moral inclusion/exclusion 121, 122, 262–265
 - of others xxxii, 260, 263
 - of women 72
 - see also* TOI (self-transcendence, openness and inclusion)
- inclusive fitness 5
- India 131
- individual virtue 58–59

- individualism 183, 230n.27
 - individualism–collectivism 64
 - individualistic culture 34
- information
 - and empathy 45, 46
 - evaluating 43
 - fMRI data 31n.5
 - internal sense 45
 - and moral emotions 222–224, 228–230, 232, 234, 237
 - and motivation 222
 - negative 43
 - and neuroscience 49
 - processing of 32, 33, 38
 - reward 38
 - social transmission of 17
- Inglehart, R. 59, 64, 195n.18
- injustice 73, 121, 153, 201, 217, 220, 267, 274
 - and anger 222
 - recognizing 275
 - reducing 263
 - seeing as misfortune 201
 - seeing in nonmoral terms 201
 - social 201, 224, 275
 - unacknowledged 221
- institutionalization 66, 67–68, 73
- instrumental consistency 126–127
- insular cortex 45
- integrative personality 93–98
 - cognitive-affective processing system 95–97
 - idiographic vs. nomothetic approaches 93–95
 - Knowledge and Appraisal Personality Architecture (KAPA) model 97
 - neo-socio-analytic model (NSAM) 97–98
 - personality architecture 97
 - three-level approach to personality systems 98
- integrity
 - behavioral 124–125, 127, 133
 - identity 123–124, 128
 - moral 165
 - restoring 166
- internal locus of control 105
- interpersonal neuroscience 34
- interruption/trying not to interrupt 153n.19
- Islam 61
- Jackson, P. 273
- JafariNaimi, N. 252n.16
- James, W. 203n.22, 219, 268n.37
- Janoff-Bulman, R. 265
- Japan
 - compared with the US 274
 - corporations in 62–63
 - Hokkaido frontier 64
 - in individualism 230
 - moral diversity patterns 63
 - Noh* theater 157
 - on shame 229, 235
 - Tale of Genji* 228
- Jaspers, Karl xxviii.28, 251n.12
- Jenkins, J. M. 226, 233n.34, 234n.35
- Johnson, D. J. 198
- Johnson, E. J. 180
- Johnson, S. L. 229
- Jones, T. M. 66
- Joseph, C. 8
- judgment–action gap xxx, 122
 - and moral identity 116–117
 - overcoming 124–125
- judgments viii, xvii, xviii, xix, xxviii, 16, 33, 233
 - biasing 43
 - conventional 63
 - emotion-laden 43
 - explicit xix
 - implicit 43
 - nonmoral xxi
 - norm xviii
 - principled xxix
 - and pure reason 40
 - self-judgment 38
 - unconscious social processes underlying 59
 - utilitarian 40
 - see also* moral judgments
- Junge, T. 291–292
- Kahneman, D. 192
- Kant, I. xxn.19, 117n.10, 145
- Keane, W. 221
- Keefer, M. W. 162
- Keltner, D. 226, 233n.34, 234n.35
- Kernes, J. L. 63
- Kesebir, S. 60
- Kierkegaard, S. xxn.19, xxviii.28, 48n.29, 118, 146n.3, 189, 216n.2, 251, 252n.14, 253n.17, 260n.25, 271
 - Kierkegaardian “leap” 251–252
- Kihlstrom, J. F. 30n.2
- Killen, M. 198, 205, 275n.46
- kin (family) selection 5–6
- kind environments 160
 - designing 167
- Kinnier, R. T. 63
- Kitayama, S. 34, 38, 62, 64, 71, 78, 131
- knowledge
 - abstract x, 267
 - and application 251
 - cultural 41
 - domain-specific 156
 - elaboration 162
 - ethical 251
 - execution 162
 - and expertise 162
 - general xi
 - identification 162
 - implicit 145n.1
 - moral 38
 - objective 251
 - planning 162
 - processing 38
 - professional 251
 - propositional 220, 221
 - real xi

- self-knowledge 37–38
- self-reflection 147–150
- semantic 37
- and skills viii, xxxi, 39, 145–176
 - application 168
- social 19
- theoretical xi
- Knowledge and Appraisal Personality Architecture (KAPA) model 97
- Kohlberg, L. xxiv, xxviii, xxxi, 60, 63, 66, 76, 93, 124, 125, 130n.27, 146n.2, 179n.1, 179, 180, 189, 194, 202–204, 206, 269, 270, 273
 - cognitive development research system 181–189, 202
 - conscious phenomenalism 185
 - critiques of system 187–188
 - responses to 188–189
 - dilemma-based assessment model 186
 - explicit judgment 184–185
 - interviews of 61
 - “From is to Ought” 182, 184
 - judgment–action gap 116–117
 - just community school 57, 69
 - and Kant 117n.10
 - “neo-Kohlberg” theory 161n.31, 188–189
 - relativism, challenging 185
 - socialization, Kohlberg’s critique of ix, xxv, 185
 - stage theory 182, 183, 185
 - system for challenging moral relativism 182–187
 - universality claims 186–187
 - see also* moral reason
- Koltko-Rivera, M. E. 262, 266
- Koole, S. L. 147n.7, 150n.14, 151n.15, 153–154, 154n.21, 265
- Kopelman, S. 91
- Korsgaard, C. M. 19
- Kramer, R. M. C. 182
- Krebs, D. L. 12, 13, 14
- Kuhl, J. 147n.7, 150n.14, 151n.15, 265
- language xiii, 9, 106, 122, 135, 164n.23, 258, 263
 - causal 31
 - and emotions 216, 219
 - ethical 69
 - everyday 145
 - German 249
 - see also* *Bildung* (education as personal formation)
 - and grammar 14
 - and human rights 134, 250, 274
 - information-sharing 19
 - of justice 134
 - module 193
 - monastic/mystical 132, 134
 - and moral formation 280–281
 - and personality 103
 - software design 236
- Lapsley, D. K. xix, xxv, 102–103, 120n.16, 276n.48, 278
- Lasky, B. xix, 102–103
- Latane, B. 148n.10
- Latin American cultures 37n.16
- leadership 73–74
- League of German Girls xiv
- Leary, M. R. 14
- Lebois, L. A. M. 30–31, 220
- Leveson, N. 67–68
- Levi, P. 291
- Lewin, K. vii, viii, x
- Lewis, G. J. 6
- listening skills 153n.19
- Little, B. R. 96, 119f
- local ecology 21
- logical entailment 125, 126
- logos xxvi
- loyalty xiv, xvi, xvii, 73, 236, 274
 - decision tree xvii
 - “double-edged sword” xvii
 - group 7, 75, 96, 101, 194, 228n.24, 264
 - party, xvi
- Luo, Q. 43
- McAdams, D. P. xxx, 89, 98–99, 101, 103–106, 258, 268n.37, 269
- McClintock, Barbara xixn.16
- McGregor, I. 96, 119f
- Malle, B. F. xviii, 197–198, 205
- manipulation of others 18
- Markett, S. 35
- Markus, H. R. 59, 71, 78, 131
- Marsella, S. xixn.17
- Marshall, S. P. 162
- Maslow, A. xvi, xxvin.27, 132, 248n.5, 262, 266
- mathematical modeling 10
- Matsuba, M. K. 91, 100
- mature moral functioning 159
- May, M. A. xxiii, xxiv, 99, 116, 269
- May, R. 48n.29
- Mayer, R. 93
- medial orbitofrontal cortex *see* mOFC (medial orbitofrontal cortex)
- medial prefrontal cortex *see* mPFC (medial prefrontal cortex)
- Meister Eckhart (mystic) 249, 250
- Menon, U., xx
- Merton, T. xiv–xv, xvi, xx, 131, 132, 134–135, 232n.33, 254, 279
 - Conjectures of a Guilty Bystander* xiii
 - Letters to a White Liberal* xiii
- Messick, D. M. 91
- Michelangelo phenomenon 70, 71, 75
- middle frontal gyri 36
- Mierop, A. 31
- Milgram, S. ix, 66, 90, 201
- Miller, D. T. xxi
- Miller, G. F. 5, 15
- mindfulness 38, 154
- mirror neurons 32
- Mischel, W. xxiv, 91, 94–99, 116, 118
- mOFC (medial orbitofrontal cortex) 36, 40

- Molaison, H. G. (H. M., epilepsy patient) 29, 30n.2 48, 49
- Moll, J. 17, 41, 44, 45, 117
- monasticism/monastic and spiritual formation 132, 134, 253–255, 271, 272–273
- Christian monasticism 275
- historical aspects 276
- see also Bildung* (education as personal formation); religion
- Montag, C. 35
- Moore, G. E. 179n.1
- moral action
- collection of capacities underlying 21
- controlled and automatic 168–169
- defining “moral” and “action” viii
- landscape of xxv–xxxii
- personality influences on 98–106
- scope of xvii–xxiii
- search for variety in xxvii–xxviii
- tensions in the field xxiii–xxv
- tied to social nature of being human 20
- treating as a design problem 167
- ultimate status of moral claims 293
- variety in 292
- see also* action; moral attentiveness; moral behavior; moral character; moral collapse; moral contempt; moral decoupling; moral development; moral dumbfounding; moral ecology; moral emotions; moral evaluation; moral exclusion; moral formation; moral identity; moral imagination; moral judgments; moral licensing; moral potency; moral pride; moral principles; moral psychology; moral reason; moral relativism; moral self-contempt; moral tightening and loosening
- moral attentiveness 155
- moral behavior viii, ix, xvii, xxiv, 8, 29, 76, 155, 162, 163, 270
- moral identity and the self 122, 124, 127, 131, 133
- moral character viii, xxiii
- moral claims, normative status xxviii
- moral cognition 203–204
- see also* cognition
- moral collapse 199
- moral decoupling 201
- moral development xv, 136, 206, 248, 260, 264, 270
- empirical description 277
- Kohlberg’s system 63, 182, 184, 185, 189, 206
- and moral reasoning 186, 206
- morality and immorality xxix, 4
- psychological approach to ix, 182
- research program 60
- stages of 63, 184, 185
- of women 188
- see also* development
- moral distancing and failure 200–202
- modifying attitudes toward responsibility 201–202
- modifying attitudes toward the victim 201
- modifying the attitude towards consequences 201
- modifying the descriptions of the action 200–201
- moral diversity across and within cultures 59–65, 78
- class differences 62
- individual variation in goals 62–63
- patterns across cultures 63–65
- political values within 60–61
- as a problem and as a solution 76
- recognition not requiring radical relativism 76
- religious influence 61–62
- structure to 75–76
- see also* moral ecology
- moral dumbfounding 191
- moral ecology xxx, 32–35, 57–86
- and application 77
- attention 71–72
- close relationships in friendship and community 69–71
- construal 72
- conventional 189
- culture and moral diversity 59–65
- as a design issue xxx, 57
- and education xxxii
- evolutionary theories 21
- individual virtue 58–59
- influencing 71–74
- interaction with moral formation 280–281
- leadership 73–74
- levels, interaction of 78
- moral exemplars and social reformers 72–73
- and moral identity 122
- motivation 72
- open questions 77–78
- at the organization level 65–69
- organizational values and goals 65
- planned, moral formation as 272–273
- processes supporting organizational corruption 66–69
- rationale for 58
- and social influence 35
- supporting domain limits in moral judgment and action 75
- see also* moral diversity
- moral emotions viii, xxxi 42–44
- anger 223–224
- and application 235–236
- as bias 42–43
- compassion 226–227
- contempt 122–130
- disgust 223
- elevation and awe 73, 231–232
- embarrassment 227–228
- and evolutionary theories 8
- gratitude 232
- guilt 225–226
- identifying 221–232
- including in models of reflective equilibrium 235
- integration into teaching of ethics 235
- learning 233–234

- not a passive process 235
- open questions 236–237
- other-condemning 222–225
- other-praising 222, 231–232
- other-suffering 222, 225–227
- pride 228–229
- questioning the cognitive/emotion distinction 44
- role in prioritizing moral thoughts, goals and action 237
- self-conscious 227–230
- shame 229–230
- see also* emotions
- moral evaluation
 - complexity of 128, 180, 246
 - and moral emotions 235
 - and moral reasoning 206
 - occurring where 205
 - programmatic approach to 258
 - social-emotional reflective equilibrium 294, 295
 - and US 2020 presidential election 235
- moral exclusion 155
- moral exemplars xv, xix, xxii–xxiii, 74
 - compassion 263
 - in computing 119n.15
 - consistency 128
 - in engineering 156, 163
 - individual variation in goals 61
 - and matched controls 258
 - and moral cognition 203
 - and moral formation 273
 - moral self, critiques 132
 - and personality 92, 93
 - and social reformers 72–73
 - sub-types 263
- moral failure 115, 200, 202
 - defining 199
 - modifying attitudes toward responsibility 201–202
 - modifying attitudes toward the victim 201
 - modifying the attitude towards consequences 201
 - modifying the descriptions of the action 200–201
 - and moral distancing 200–202
 - paths to 199–202
 - understanding/learning from 169
- moral formation viii, xxxi–xxxii, 44–46, 246–290
 - and application 278–279
 - Bildung* 249–250
 - critique from traditions 258–259, 271–272, 275–276
 - as cultural conservation and critique 274
 - dimensional landscape for 295–297
 - direction of change in 269–271
 - domain limited psychological models 150
 - educative aspects 272–276
 - empathy 45–46
 - evaluative–self transformative aspects xxxii, 248, 256–259
 - existential appropriation 251–253
 - facilitative/educative aspects 248, 256
 - goal-directed 267
 - Greek *paideia* 250, 251
 - humanistic/experiential traditions xi
 - inclusion/exclusion 262–265
 - interaction with moral ecology 280–281
 - and language 280–281
 - monastic/spiritual formation 253–255
 - and moral exemplars 273
 - and moral identity 246–247
 - and moral luck 275
 - normative claims 277
 - normative/self-transcending aspects xxxii, 248, 256, 259–272, 277–278
 - occurring across multiple domains/dimensions 278
 - open questions 279–281
 - openness to change 262
 - personal, interpersonal and cultural critique, concern for 276–277
 - as planned moral ecology 272–273
 - psychological models 150
 - Roman *humanitas* 250
 - rules vs. goals 265
 - and self-transcendence 254, 260–262, 266, 270, 272, 278, 279–280
 - as social influence 273–274
 - specific commitments and projects 266–268
 - storytelling 257–258
 - and taking moral action 247
 - terminology 246
 - themes and questions from the traditions 255–256
 - three-dimensional space (self-transcendence, openness and inclusion) 268–269
 - traditional accounts 249–256, 280
 - value dimensions of normative formation 259–268
 - value space of 46, 260
- moral function, and neural systems 30–31
- moral identity viii, xxviii, 155
 - and application 134–135
 - behavioral integrity 124–125
 - and community 70
 - consistency 125–128, 129*t*
 - defining 117–122
 - formation xxv
 - forms of 123–125
 - holding each other accountable 128, 130
 - identity integrity 123–124
 - and moral ecology 122
 - and moral failure 200
 - and moral formation 246–247
 - open questions 135–136
 - and personality 91
 - and the self xxx, 37–38, 115–144
 - “What is the situation like?” question 115
 - “Who is a person like me?” question 115, 119
- moral imagination 128, 155–156
- moral intensity 66
- moral judgments ix, xvii, xix, xx, xxvii, xxviii, xxx, 3, 8, 14, 15, 16n.13 40–42, 47, 63, 269
 - and decision-making 39

- moral judgments (*cont'd*)
 - and emotion centers 39–40
 - and moral identity 115
 - political values 60
 - relevant skills 146
 - supporting domain limits in 75
 - types of xviii, 197–198
 - see also* explicit judgments; judgments
- moral licensing 126
- moral luck 275
- moral norms 14–15
- moral philosophy xxi, 179
- moral potency 155
- moral pride 228
- moral principles 75, 96, 128, 182n.4 205
 - selective use of 202
- moral psychology xxvii
 - and developmental psychology ix, 181
 - evolutionary theories 22
 - as a fragmented field x–xi
 - framework for moral evaluation and commitment 294–295
 - historical perspective viii–x
 - instrumentalist approach in 205
 - interdisciplinary conversation xviii
 - isolation from study of real moral action x
 - literature 91
 - and personality 92
 - philosophical xxi–xxii
 - Piagetian 146n.2
 - psychological xxi–xxii
 - as psychology of evil 75
 - and reason/reason-emotion tension 180, 205–206
 - research 184–185
 - theological xxii
 - theoretical framework for 293–297
 - see also* psychology
- moral reason viii, 179–214
 - and application 204–205
 - clarity of meaning, need for 203
 - implicit cognition and two-process models 189–198
 - and moral evaluation 206
 - naturalistic moral cognition 198–199
 - open questions 205–206
 - paths to moral failure 199–202
 - type of 180–181
 - unconscious 203
 - as a useful guide and source of bias 204
 - see also* Kohlberg, L.; reason
- moral relativism
 - critique 182–184
 - Kohlberg's system for challenging 182–187
- moral self 75, 106, 117, 118, 119n.14, 120, 122–124
 - active in maintaining web of commitment and action 135–136
 - critique 130–132
 - moral failure integrated into 136
 - multidimensional 133, 135
 - omnipotent 134
 - see also* moral identity; moral self-concept; self
- moral self-concept 37, 118, 120
 - influencing action 122–130
 - see also* moral identity; moral self; self
- moral tightening and loosening 121–122
- morality
 - community-based 70
 - content xx
 - conventional 61, 187
 - dark side 20
 - as a design issue 76
 - and emotions xxxi, 220–221
 - and ethics 186
 - evolutionary accounts 7, 12–15, 21, 22
 - expertise model 157
 - function xx
 - heteronomous 183
 - and immorality 22
 - importance of integrating into the self 132–133
 - integration into the personality 108
 - post-conventional 187, 188
 - pre-conventional 188
 - relationship to life themes/projects/stories 108
 - and religion xvi
 - restrictive definitions xix
 - separation of skills and knowledge from 166
 - see also* emotions; ethics; evolutionary theories; expertise; moral action; religion
- Mormons 61
- Mothers of the Plaza de Mayo xii, xv, xx, 132, 150, 152, 168
- motivation xiv, xxvi, 9, 155, 259
 - action 216
 - and emotions 127, 221, 222, 230, 231, 234
 - to ethical action 252
 - human distinctiveness 15
 - and information 222
 - moral xxv, xxxiii, 228, 260, 282
 - and moral ecology 71, 72
 - and moral emotions 223, 224, 227
 - and moral identity 117, 121, 124, 125, 127, 130n.27, 135
 - and moral reason 183, 198
 - and personality 100, 104
 - pro-social behavior 221
 - “selfish” 131
 - self-transcendence 266
 - storytelling 257–258
 - and withdrawal 230
- Moulton, C.-a. E. 165
- mPFC (medial prefrontal cortex) 32, 33, 43
 - resting metabolic rate 37–38
 - see also* brain; neuroscience
- MRI (magnetic resonance imaging)
 - structural scans 36
 - see also* fMRI (functional magnetic resonance imaging)
- Mullins, D. A. 63
- multidimensionality of the self xxx, 118–120, 133
- Mumford, M. D. 91
- Munson, R. (pseudonym in volunteerism study), and family xii, xiv, xv, xx, 20, 21, 132, 247, 263–264, 278, 279
- Muslims 61

- Muthukrishna, M. 63
mutualism 9
- narrative consistency 127
narrative identity 106
narratives 15, 35, 98, 257–258
 autobiographical 199
 master 274
 moral 158
Narvaez, D. xxv, 147, 154, 155, 157, 158,
 159, 162, 188, 189, 202n.21, 248n.5
Nathan, L. 252n.16
natural history xix, xxi
natural selection 4, 9
naturalistic fallacy 179n.1, 184, 206
naturalistic moral cognition 198–199
nature vs. nurture xxiv
Nazism xiv, xx
Neal, D. T. 164–165
“neo-Kohlberg” theory 161n.31, 188–189
neo-socio-analytic model (NSAM) 97–98
neural unification 34
neuro-essentialism 31
neurons 32
neuroplasticity 39
neuro-realism 31
neuroscience viii, xxix–xxx, 29–56
 and application 48–49
 causal models 47, 49
 connection of social and neural processes 49
 cultural, group and individual influence on
 brain processes 33–34
 imitation and mirror neurons 32
 interpersonal 34
 moral ecology 32–35
 neural systems and moral function 30–31
 open questions 49
 thinking about, and with, others 32–33
 using 30–31
 see also brain
neuro-synchrony 46
neuroticism–emotional stability personality
 dimension 36, 100–101
Nisan, M. 126
Nisbett, R. E. 62
Nissenbaum, H. 67
nomoethic approaches, vs. idiographic
 approaches 93–95, 98
nonconscious processing 189
non-egocentric self 253, 254
non-humans/primates
 imitation and mirror neurons 32
 interdependence among 12–13
 macaque monkeys 32
 proto-morality 4, 21
 self-consciousness 18
 social consciousness 18
 universalized normative evaluations 19
norm judgments xvii, 197
normative claims 270, 277, 292, 293
normative moral formation xxxii, 256, 263,
 264, 268, 269, 271, 273, 274
normative “ought” statements 179, 182, 206
normative philosophy 183, 184
normativity viii, xxvi–xxvii, xxviii, xxxi, 118,
 179, 181
 in action xxvii, 268
 cognitive models 199
 ethics 270
 evaluative vs. normative 259
 and objectivity 251n.11
 rational vs. moral normative approaches 183
 and self-transcendence 277
 social appraisals 219n.6
 universalized evaluations 19
 values xxxii, 248, 277, 279
 see also moral formation; norm judgments;
 normative claims; normative moral
 formation; normative “ought” statements;
 normative philosophy
novices 160
NSAM (neo-socio-analytic model) 97–98
Nucci, L. P. 130
nucleus accumbens 43
nurturance 9
Nussbaum, M. 250
- Oatley, K. 226, 233n.34, 234n.35
obedience to authority, “shock studies” ix
OFC (orbito-frontal cortex) 38, 40, 41
Oliner, L. P. 92, 226–227
Oliner, S. P. 92, 226–227
Omoto, A. M. 121
openness
 to change xxxii, 261, 262
 dispositional moral traits 101
 openness–closedness 101
 see also TOI (self-transcendence, openness and
 inclusion)
Opotow, S. 121, 263
optimism and positivity 105
orbito-frontal cortex *see* OFC (orbito-frontal
 cortex)
organizational corruption, processes supporting
 institutionalization 67–68
 rationalization 66–67
 socialization 68–69
organizations xii, xxii, 20, 77
 charitable 21
 codes of ethics 68
 control-oriented 73
 customer-driven 65
 and cynicism 105
 diversity programs 68
 ethical action 71
 finance-driven 65
 hierarchical 73
 influence of moral action on xviii
 institutionalization 67
 leadership in 73
 and local units 69, 71
 moral ecology at level of 65–69
 organizational climate 58
 organizational culture 58

- organizations (*cont'd*)
 - organizational ethics as a design problem 71–72
 - quality-driven 65
 - socialization in 68, 273
 - unethical conduct 65
 - values and goals 65
 - volunteer 75
 - see also* organizational corruption, processes supporting
- Orom, H. 90
- ought-based reasoning viii, xx, xxii, 16, 184, 258
 - and Freud xxn.20
 - normative “ought” statements 179, 182, 206
- Over, D. 181
- oxytocin 31

- Pals, J. L. xxx, 89, 98–99, 103
- paradox of control 68, 70
- parallelism 183, 184
- parietal lobes 46
- Park, R. 58
- Paunov, Y. ix–x
- PC (posterior cingulate) 36
- Penner, L. A. 90
- Peper, M. 216n.1, 219n.6
- perception–judgment–action 159
- personality viii, xxx, 35–36, 89–114
 - as adaptive 89, 90–92
 - and application 107–108
 - architecture 97
 - and behavior xxiv, 91, 94–96, 148, 269
 - characteristic adaptations 35, 104–106
 - coherence 127n.23
 - cynicism 105
 - dispositional moral traits 99–103
 - five-factor model *see* Big Five structure of personality
 - generativity 105–106
 - genetically determined physiological theories xxiv
 - influences on moral action 98–106
 - as integrative 89, 90, 93–98
 - internal locus of control and self-efficacy 105
 - and moral exemplars 92, 93
 - narrative identity 106
 - no singular moral personality 107
 - open questions 108
 - optimism and positivity 105
 - paradox *see* personality paradox
 - as pluralistic xxx, 89, 90, 92–93
 - pro-social 104–105
 - resilient 91
 - signatures *see* personality signatures
 - vs. situation xxiii–xxiv
 - three-level approach to systems of 98
- personality paradox 96, 97
- personality signatures 95, 96
- Pessoa, L. 44
- Peter-Hagene, L. C. 223
- Peterson, C. 102
- PFC (prefrontal cortex) 43
- philosophical moral psychology xviii.12, xxi–xxii
- philosophy
 - reflective equilibrium, philosophical roots of term 281, 294
 - and theology xxi–xxiii
- physical ecology 21, 22
- Piaget, J. 146n.2, 185, 186
- Pilcher, J. J. 10, 11, 12
- Piliavin, J. A. 122, 264
- Pincoffs, E. 186
- pISTS (posterior left superior temporal sulci) 36
- Pittman, T. S. 15
- Pizarro, D. A. 202
- Plato 89, 102
- pluralism xxx, 92–93, 188
- political values 60–61
- posterior cingulate *see* PC (posterior cingulate)
- posterior left superior temporal sulci *see* pISTS (posterior left superior temporal sulci)
- posterior temporo-parietal junction 46
- Postle, B. R. 31n.5
- Power, F. C. 57, 69
- Pratt, M. W. 122, 264
- preconscious processes 181
- prefrontal cortex *see* PFC (pre-frontal cortex)
- prejudice research ix
- pride 228–229
- primates *see* non-humans/primates
- Pritchard, M. 156
- processes vii, viii, xi, xxv, xxvi, xxvii, xxviii, xxix, xxxi, xxxii 31, 46, 60, 65, 69, 73, 78
 - biological 29
 - cognitive and emotional xxxi
 - cultural 7, 17, 30, 59
 - decision 71
 - developmental 7
 - emotion-reason 47, 49
 - of empathy 33
 - evolutionary 7, 10, 15, 76
 - friendship and community 69–70
 - group 13, 77
 - interweaving of 59
 - of moral ecology 74, 75
 - moral psychology 75
 - motor 37
 - neural/neurological 29, 32, 33–34, 36, 39, 47, 48, 49
 - organizational 66–69, 71
 - physiological 35, 48
 - psychological xv, xxx, 32, 36, 38, 44, 49, 50, 59
 - self-regulation 17
 - social 49, 59
 - supporting organizational corruption 66–69
 - time 30
- Profet, M. 12
- propositional knowledge 220, 221
- pro-social behavior ix, 13, 21, 22, 90, 104, 109, 221, 232
- pro-social personality 104–105
- Protagoras 89, 94, 106

- proto-morality of non-humans *see* non-humans/
primates
- Przybylski, A. K. 130, 254
- psychology
clinical 150
developmental 40
of evil 75
evolutionary 10
psychological moral psychology xvii, xxi
psychological selfishness 9
psychological themes xxvii–xxviii
sub-fields 15
see also moral psychology
- psychopaths 166
- PTSD (post-traumatic stress disorder) 42n.24
- punishment 19, 67, 95
altruistic 3, 7, 8, 22, 222, 224
anticipatory 3, 13, 14
justifiable 43
moral violations 224
and reward 197
vicarious 10
- quandary ethics 186
- Quinn, J. M. 164–165
- Quirin, M. 150n.14
- racial hatred xiii, 107
- Raffensperger, B. xiv, xvi
- rationalism xxi, xxxiv
- rationality 21, 184, 203, 204
conscious 187
and culture 20
ecological 193, 202
and emotions 17, 47n.28, 235, 236
ethical 205
evolutionary theories 21
instrumental nature of 181
models 180, 181
moral 193
rational regulation 49
scientific 42n.22
self-regulation of 47n.28
type 1 194
see also reason
- rationalization 66–67, 73
- Rawls, J. 202
- reason
autobiographical reasoning 198
balancing with social/emotional reward
40–41
vs. bias xxiv–xxv
blending with emotion 17, 41, 49
content-specific reasoning 16n.13
going awry 199–202
instrumentalist 180, 181
moral judgment and reasoning 14
planning for cognitive-emotional
effectiveness 41–42
prescription of correct reasoning vs.
description of how people reason 181
pure 40
reason/emotion dichotomy 47, 220
symbolizing and reasoning 16
theoretical reasoning 180
see also moral reason; rationality
- reciprocity 9
- reconciliation model 124
- reflective equilibrium
including emotion in models of 235
and moral failure 202
philosophical roots of term 281, 294
see also social-emotional reflective equilibrium
(SERE)
- regulation
emotional *see* emotional regulation
of habits 164–166
rational 49
see also self-regulation
- relativism xxvii, xxviii, xxxi, 76, 125n.20, 179,
204, 221
cultural 75, 76, 184, 185, 188
of Durkheim 69n.10
Kohlberg's challenge of 185
radical 76
rejecting xxxiv
simplistic xxviii, 276
see also moral relativism
- religion xii
Abrahamic traditions 48n.30
Advaita Vedanta Hinduism 254
Buddhism 226n.20, 253, 254, 258
conversion 150
Hindu culture 216n.2
monastic/mystical traditions 132, 134,
253–255, 271
moral thinking of Jewish people and Christian
Protestants in the US 61
and morality xvi
religious influence 61–62
status 19
Sufism 254
and theological moral psychology xxii
viewing person as an image of God 249
see also *Bildung* (education as personal
formation); theology
- resilience 91, 169, 226n.20
- responsibility 130, 200
collective 57
denial of 67
diffusion of 67, 201
displacement of 67, 201
emotional 130, 136, 254, 257
individual 64
modifying attitudes to 201–202
moral 67
and self-within-context 124
social 104
- Rest, J. R. xxv, 188
- Restall, G. 126n.21
- restorative justice xiii, xv–xvi, 133, 230n.28
and personality 107–108

- Reuter, M. 35
 Reyes, M. R. 233–234
 Reynolds, S. J. 66, 68, 155
 Richerson, P. J. 7
 “right thing” desire to do xx
 Rivers, S. E. 233–234
 Roberts, B. W. 92, 97–98, 103, 104, 148, 149f, 150, 258
 Roman Catholic Church 92
 Rothman, A. J. 164
 Rudd, A. 125n.20, 275n.45
 Rugh, R. 34
 Rule, J. T. 122, 264
 Rule of Benedict 232n.33, 273
 Rusbult, C. E. 70
 Ryan, R. M. 130, 254
- Salerno, J. M. 223
 Sapthiang, S. 154
 Satyal, M. K. 34
 Saucier, G. 101
 Sayre-McCord, G. 179n.1
 Schein, C. 193n.15
 Schelling, F. W. J. 189
 schema theory 162
 Scherer, K. R. 216n.1, 219n.6, 224
 Schmitt, D. P. 10, 11, 12
 Scholl, S. M. xiv, xv, xx, 267, 291
 Schön, D. A. 165
 Schwartz, S. H. 59, 65, 93, 123, 195, 263, 269, 270, 296
 circle of value regions 63f, 196f
 Schwitzgebel, E. 128
 scope of moral action xvii–xxiii
 exploration vs. definition xviii–xxi
 philosophical moral psychology xxi–xxii, xviii.12
 philosophy and theology xxi–xxiii
 psychological moral psychology xvii, xxi
 theological moral psychology xxii
 second order desires/volitions 17, 44
 Selective Table of Definitions of Moral Psychology xviii–xix, xxiv–xxv
 self
 construction and maintenance of 60
 critiques of moral self 130–132
 “deep social roots” of 122
 different roles of 135
 as domain specific 121–122
 ideal 70, 122
 importance of integrating morality into 132–133
 life-long transformation of 131
 moral *see* moral self
 and moral identity xxx, 37–38, 115–144
 moral self-contempt influencing action 122–130
 as multidimensional xxx, 118–120, 133
 self-within-context 124–125
 situational variation in 121
 as variable over time 120
 varieties of 117–122
 self-actualization 19, 132, 266
 self-awareness 18
 self-concept xv, 12, 73, 103, 116, 119, 124–125, 267
 moral 37, 118, 120, 124
 self-consciousness/self-conscious emotions 18, 222, 227–230
 neuroscience 29, 37, 38
 self-control system 17
 self-regulation 150, 151–152
 situational self-control 42
 self-efficacy 105
 self-enhancement 261
 self-ideal 122, 264
 self-interest xiii, xx, xxi, xxvii, 9, 40, 132
 economic decision-making 40
 selfishness 4, 7, 8, 15
 vs. altruism 9–10
 self-knowledge 37–38, 41
 self-maintenance system 151, 152–153
 self-reflection xxx, 128
 exit points 148–149
 flexibility of as a guide for moral action 133–134
 skills and knowledge 147–150
 self-regulation xxii, 90, 95, 150–155
 and cognitive control 17
 emotional 151
 general skills 146
 and habits 165
 and moral formation 44
 self-control system 150, 151–152, 155
 self-maintenance system 151, 152–153
 self-reports 34, 102, 122, 191n.12 195, 225
 of autobiographical memory 281
 of moral behavior 270
 studies of guilt 225–226
 of traits 264
 self-transcendence xvi, xxxii, 19–20, 64, 132
 and existentialism xxvi
 and hierarchy of needs xvi
 and moral formation 254, 260–262, 266, 270, 272, 278, 279–280
 see also TOI (self-transcendence, openness and inclusion)
 self-transformation 250, 253, 255
 Seligman, M. E. P. 102
 semantic knowledge 37
 sensorimotor cortico-basal ganglia, habit-based 39
 Serafeim, G. 67
 SERE *see* social-emotional reflective equilibrium (SERE)
 sexual selection 5, 7
 Shakespeare, W. 234n.36
 shame 229–230
 Sheeran, P. 164
 Shein, C. xxn.21
 Shenhav, A. 40, 41
 Shoda, Y. 95, 97
 Shonin, E. 154
 Shweder, R. A. xx, xxviii, 197, 219n.6
 Sinnott-Armstrong, W. ix, xix, xx, xxi, 47n.26
 situational influence 58
 situational self-control 42

- skills
 defining what is moral in 166
 domain limited 167
 domain specific 146, 156, 166, 168
 emotional self-regulation 151
 expert performance 157–163
 general 146–156
 going awry 166
 involved in moral action 168
 and knowledge viii, xxxi, 39, 145–176
 moral attentiveness 155
 moral imagination 155–156
 self-control system 150, 151–152
 self-maintenance system 151, 152–153
 self-reflection 147–150
 self-regulation 150–155
 skill development and moral action 280
 supporting moral action as domain-limited 167
- Smith, Adam 189
- Snyder, M. 121
- Sober, E. 6
- social brain 3–4
- social consciousness 18
- social consistency 126
- Social Darwinism 4n.3
- social ecology 21, 22, 58
- social identity theory 74
- social influence, moral formation as 273–274
- social intuitionist model 126
- social reformers 72–73
- social status 20, 131
- social-emotional reflective equilibrium
 (SERE) xxxi, xxxii, 128, 180, 271, 281,
 294–295
 establishing 236–237
 and moral emotion 235, 236
 and moral reason 202, 204, 206
- socialization 66, 68–69, 71, 73, 273
 Kohlberg's critique of ix, xxv, 185
- socio-technical system 58
- Socrates 89, 94, 106, 251, 271
- Solomon, R. C. 216n.2, 217n.3
- sophomore retrogression 182
- Southwest Airlines 72
- Soutschek, A. 46
- Spencer, H. 4
- stability
 behavior 96n.11
 vs. change 46, 148
 cognitive-affective processing system 95–97
 vs. development xxiv
 emotional 100–101
 personality meta-trait of 35
- Stanovich, K. E. 190
- Steinberg, L. 41
- Stellar, J. E. 62
- storytelling 15, 257–258
- strategic interaction 13
- Strauman, T. J. 45
- Stürmer, S. 121
- Sufism 254
- symbolizing and reasoning 16
- Tamir, D. I. 33
- Tangney, J. P. 225n.16, 230n.28
- Tannenbaum, D. 71, 202, 205
- Taylor, S. E. xxiv
- Tell, C. E. 195
- temporal lobes 46
- Thalhammer, K. xii, xiii, 156
- theology xxi–xxiii, 48n.29
 theological anthropology 249
 theological moral psychology xxii
see also religion
- theory vii, ix, 117, 124, 162, 184, 297
 as abstract knowledge x, xi
 decision theory 180
 developmental 161
 and emotions 215, 217n.3, 219
 evolutionary *see* evolutionary theories
 integrative theory-making 249
 of moral action *see* moral action
 moral formation *see* moral formation
 moral foundations *see* moral foundations
 moral reasoning 187, 195, 196–197
 and neuroscience imaging 33, 50
 of personality 89, 95
 and practice/practical action x, xi, 160
 psychological xxii, xxiv, 11, 30–31, 132, 194,
 202n.21 292
- scientific xix
- theoretical coherence 10
- theoretical reasoning 180
- unifying xxv
 of value xxv, 79, 195
see also theory of mind (ToM); virtue theory
- theory of mind (ToM) 13, 15n.12, 18, 19, 21,
 46, 155, 188, 233
- Thích Nhất Hạnh 254
- Thoma, S. 188
- Thornton, M. A. 33
- Tiberius, V. xxi, xviii.12
- Tillich, P. xx, 48n.29
- time perspective 16
- TOI (self-transcendence, openness and inclusion)
 xxxii, 268–269, 277, 278, 296, 297
- Tönnies, F. 60–61
- Tooby, J. 193
- Tracy, J. L. 232n.33, 254
- training systems xi–xii, 42
- traits
 character as trait approach 89, 90
 character-relevant 102
 defining 35
 mapping virtues as 101–103
 personality systems 98
 psychological status 103
 simple 116
 stability, personality meta-trait of 35
 trait-level description of personality 36
see also dispositional moral traits; personality
- transcendence 296
see also self-transcendence; TOI (self-
 transcendence, openness and inclusion)
 “tree switching” xviii.15

- Treviño, L. K. 66, 68
tribalism, human 18
Tripathi, R. 94
Trump, D. xiii, xiv
 accusations of voter fraud xx, 235–236
Tse, P. U. 14
Tufts, J. H. 183
Tversky, A. 192
two-process models 218n.5
 complexities 192
 content (type 1 processes) 192–197
 human values as 195–196
 moral foundations as 193–195
 moral modules as 193
 relational models as 196–197
 critiques 197–198
 defining 190–192
 five foundations 194
 and implicit cognition 189–198
- Uhlmann, E. L. 202
ultimate concern 252, 255, 262, 270, 271,
 278–281, 294
ultimate values xvi, xxvi, xxvii, 260, 262, 277
“ultimate verities” xvi, xxvi, 293, 294
unconscious, the 189
United States
 civil rights movement xiii, 266n.32
 compared with Japan 274
 Federal Sentencing Guidelines for
 Organizations 67
 gay rights movement 274
 human rights in xiii
 moral thinking of Jewish people and Christian
 Protestants in 61
 openness to change 262
 organizational ethics programs 67
 peace movement (1960s) 134
 political climate 61
 presidential election of 2020 72–73
 shame in 229–230
 social service in 92, 281
 southern 64
 spiritual formation/direction in 254
- Vaillant, G. E. 248n.5
Valadez, S. 266
Vallor, S. xixn16
value internalization 68
value space of moral formation 46, 260
values
 agentic 270
 communal 270
 cross-cultural value difference 65, 93
 engineering 65
 organizational 65
 political 60–61
 religious 61
 two-process models 195–196
 ultimate *see* ultimate values
 value pluralism 103
values practices 68
- Van der Toorn, J. ix–x
Van Gordon, W. 154
Van Langevelde, A. xiii–xiv, xvi, 236
Van Nunspeet, F. 38n.19, 46
Van Oudenhoven, J. P. 102
vенеer theory 15
ventro-mPFC *see* vmPFC (ventro-mPFC)
Vianello, M. 73
victim-blaming 201
Virgin Atlantic 72
virtue ethics 160n.29 160n.29, 231n.30, 253
virtue/virtues xv, xxiii, 5
 Aristotle on 102
 beyond virtues vs. the situation 58–59
 and expertise 157
 heroic 92n.5
 individual 58–59
 mapping as traits 101–103
 philosophical virtue theory 94
 psychological status 103
 theories and moral exemplars xxii–xxiii
 and the virtuous xxx, 89, 106–107
 see also virtue ethics
vmPFC (ventro-mPFC) 38, 43
volunteerism xii, xv, 69, 75, 91, 106, 150, 264
 see also Munson, R. (pseudonym in
 volunteerism study), and family
Von Humboldt, A. xix
- Walker, L. J. ix, xvi, xxiv, 62, 99, 106, 120n.16,
 124, 127, 130, 132, 134, 258, 261, 266,
 270, 277, 296n.2
Waytz, A. 18
Weaver, G. R. 66, 68
Weber, E. U. 180
Weber, J. M. 91, 92
Weidman, A. C. 232n.33, 254
Weinstein, D. 4n2
Weinstein, N. 130, 136, 254
WEIRD (Western, educated, industrialized, rich
 democratic) societies 62, 64
Welzel, C. 59, 64, 195n.18
Westen, D. 43
Western culture 33–34
White Rose resistance movement xiv, 267
Whitehouse, H. 63
wicked environments 160
willpower 15, 17, 151, 152
Wilson, D. S. 6
Wong, D. B. 220n.6
Wood, D. 97, 98, 104
Wood, W. 164–165
World Values Survey 65
Wright, A. G. C. 35
Wright, J. C. 95
Wrzus, C. 148, 149f, 150, 258
- Yalom, I. 48n.29
- Zagzebski, L. 231n.30
Zeami 157–158
Zhu, Y. 33