

Moral Psychology

Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development

edited by **Walter Sinnott-Armstrong**

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

© 2008 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please e-mail special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Stone Sans and Stone Serif by SNP Best-set Typesetter Ltd., Hong Kong and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Moral psychology / edited by Walter Sinnott-Armstrong.

v. cm.

"A Bradford Book."

Includes bibliographical references and index.

Contents: v. 1. The evolution of morality: adaptations and innateness—v. 2. The cognitive science of morality: intuition and diversity—v. 3. The neuroscience of morality: emotion, disease, and development.

ISBN 978-0-262-19561-4 (vol. 1: hardcover : alk. paper)—ISBN 978-0-262-69354-7 (vol. 1: pbk. : alk. paper)—ISBN 978-0-262-19569-0 (vol. 2: hardcover : alk. paper)—ISBN 978-0-262-69357-8 (vol. 2: pbk. : alk. paper)—ISBN 978-0-262-19564-5 (vol. 3: hardcover : alk. paper)—ISBN 978-0-262-69355-4 (vol. 3: pbk. : alk. paper)

1. Ethics. 2. Psychology and philosophy. 3. Neurosciences. I. Sinnott-Armstrong, Walter, 1955–

BJ45.M66 2007

170—dc22

2006035509

10 9 8 7 6 5 4 3 2 1

This volume is dedicated to Mike Gazzaniga and Scott Grafton for their patience, encouragement, support, stimulation, inspiration, and friendliness while helping this neophyte learn about the brain.

Contents

Acknowledgments xi

Introduction xiii

Walter Sinnott-Armstrong

1 | The Cognitive Neuroscience of Moral Emotions 1

Jorge Moll, Ricardo de Oliveira-Souza, Roland Zahn, and Jordan Grafman

1.1 | Processes and Moral Emotions 19

William D. Casebeer

1.2 | Morality, Inhibition, and Propositional Content 25

Catherine A. Hynes

1.3 | Response to Casebeer and Hynes 31

Jorge Moll, Mirella L. M. F. Paiva, Roland Zahn, and Jordan Grafman

2 | The Secret Joke of Kant's Soul 35

Joshua D. Greene

2.1 | Moral Cognition and Computational Theory 81

John Mikhail

2.2 | Toward a Sentimentalist Deontology 93

Mark Timmons

2.3 | Reply to Mikhail and Timmons 105

Joshua D. Greene

3 | Without Morals: The Cognitive Neuroscience of Criminal Psychopaths 119

Kent A. Kiehl

3.1 | The Antisocials Amid Us 151

Ricardo de Oliveira-Souza, Fátima Azavedo Ignácio, and Jorge Moll

3.2 Impaired Moral Reasoning in Psychopaths? Response to Kent Kiehl	159
Jana Schaich Borg	
3.3 A Reply to de Oliveira-Souza, Ignácio and Moll, and Schaich Borg	165
Kent A. Kiehl	
4 Internalism and the Evidence from Psychopaths and “Acquired Sociopaths”	173
Jeanette Kennett and Cordelia Fine	
4.1 Internalism and the Evidence from Pathology	191
Adina L. Roskies	
4.2 The Truth about Internalism	207
Michael Smith	
4.3 Could There Be an Empirical Test for Internalism?	217
Jeanette Kennett and Cordelia Fine	
5 Varieties of Moral Agency: Lessons from Autism (and Psychopathy)	227
Victoria McGeer	
5.1 Reasons, Reverence, and Value	259
Jeanette Kennett	
5.2 The Will to Conform	265
Heidi Maibom	
5.3 Autism, Morality, and Empathy	273
Frédérique de Vignemont and Uta Frith	
5.4 The Makings of a Moral Sensibility: Replies to Commentaries	281
Victoria McGeer	
6 Morality and Its Development	297
Jerome Kagan	
6.1 Morality, Culture, and the Brain: What Changes and What Stays the Same	313
Nathan A. Fox and Melanie Killen	
6.2 The Fabric of Our Moral Lives: A Comment on Kagan	317
Paul J. Whalen	

6.3 Reply to Fox and Killen and Whalen	321
Jerome Kagan	
7 Adolescent Moral Reasoning: The Integration of Emotion and Cognition	323
Abigail A. Baird	
7.1 Integrative Mechanisms and Implicit Moral Reasoning in Adolescence	343
Daniel K. Lapsley	
7.2 Can Baird's View of Adolescent Morality Inform Adolescent Criminal Justice Policy?	351
Katrina L. Sifferd	
7.3 Reply to Sifferd and Lapsley	361
Abigail A. Baird	
8 What Neuroscience Can (and Cannot) Contribute to Metaethics	371
Richard Joyce	
8.1 Moral Rationalism and Empirical Immunity	395
Shaun Nichols	
8.2 Hedonic Reasons as Ultimately Justifying and the Relevance of Neuroscience	409
Leonard D. Katz	
8.3 Response to Nichols and Katz	419
Richard Joyce	
References	427
Contributors	483
Index to Volume 1	485
Index to Volume 2	513
Index to Volume 3	543

Acknowledgments

Many people deserve my thanks for assistance as these volumes grew. For financial support of the conference that sowed the seeds for this project, I am grateful to several institutions at Dartmouth College, including the Leslie Center for the Humanities, the Dickey Center for International Understanding, the Master of Arts in Liberal Studies program, the Dean of the Faculty, the Department of Psychological and Brain Sciences, the Social Brain Science Project, the Cognitive Neuroscience Center, the Department of Philosophy, and the Institute for Applied and Professional Ethics. For help in making these essays more accessible to students, I thank Cate Birtley, Cole Entress, and Ben Shear, Dartmouth students who served as my Presidential Scholars. I also greatly appreciate the devotion and editorial skills of Kier Olsen DeVries, who worked long and hard to put the essays from these diverse authors into a single form for the publisher. Tom Stone, my editor at MIT Press, also has my gratitude for his spirited encouragement. Last but not least, I thank my family—Liz, Miranda, and Nick—for their love and patience while I spent many nights and weekends on this project.

Introduction

Walter Sinnott-Armstrong

Moral psychology is old. In ancient times, Plato and Aristotle addressed many of the issues that still occupy moral psychologists, such as the conditions of responsibility and the roles of reason, emotion, and culture in moral judgments. In early modern Europe, Hume, Kant, and Mill continued these themes. Hume even used methods that anticipate some contemporary cognitive psychology. Debates about the evolution of moral beliefs, emotions, and actions have been raging since Darwin.

In contrast, the neuroscience of moral belief is brand new. Brain lesions have been studied at least since Phineas Gage, but such lesions are not numerous, controlled, or focal enough to support precise conclusions. Only within the last few decades have noninvasive techniques made it possible to get solid information about how our brains make up our minds.

Brain science concentrated at first on simple mental events and gradually gained the capacity to investigate more and more complex processes. Neuroscientific studies of moral beliefs, emotions, and decisions were not possible until the 1990s. The first brain imaging studies of moral judgments were reported as recently as 2001. The neuroscience of morality is a mere baby.

This baby is growing fast. Today many labs all over the world are planning or executing studies of the neural bases of moral judgment. The chapters in this volume sample the best work in this emerging field. They also display the variety of approaches, including functional imaging, lesion studies, abnormal psychology, and developmental neuroscience.

Some of the earliest brain imaging studies of moral judgment were performed by Jorge Moll in Brazil, who is also the lead author in the opening chapter of this volume. Moll and his colleagues Roland Zahn, Ricardo de Oliveira-Souza, and Jordan Grafman build on previous functional imaging studies as well as clinical evidence to construct a general theory of moral

emotions, including guilt, shame, embarrassment, pride, indignation, anger, contempt, pity, compassion, awe, elevation, and gratitude. This wide array of emotions is tied to certain brain regions and is unified by a certain structure. According to Moll and his colleagues, moral emotions have shared and distinctive phenomenological features that are explained by their representational components, aspects, and functions. This representational account of moral emotion raises profound challenges to the traditional dichotomy between emotion and cognition.

Bill Casebeer comments that the representational theory of Moll and his colleagues inevitably depends on substantive moral assumptions and needs to be integrated with both process views of moral emotions and with knowledge of neural reward-processing mechanisms. In her comment, Catherine Hynes argues that Moll et al. need to define which emotions are moral and that an adequate definition will show why moral emotions need to be understood in terms of inhibitory and regulatory processes as well as the propositional content of moral judgments. Moll et al. reply by clarifying their definition of morality and by showing how their representational view might be related to a process view of moral emotions.

Another early explorer, Joshua Greene, came to moral neuroscience through philosophy. In his contribution to this volume, Greene draws philosophical lessons from a broad base of empirical research in moral neuroscience and psychology, including work by Jonathan Baron and Jonathan Haidt.¹ Greene argues that deontological moral judgments and theories, which Immanuel Kant claimed to be grounded in pure reason, are actually moral rationalizations driven by emotional responses.² Consequentialist moral judgments and theories, in contrast, are more likely to be cognitive and to involve genuine moral reasoning. Greene claims that if these empirical claims are true, they cast doubt on deontology as a moral philosophy.

In response, John Mikhail shows how his computational theory accounts for Greene's data without giving up on deontological principles based on moral reasoning.³ Mark Timmons then responds to Greene's four main arguments by keeping deontology but giving up the rationalist assumption that morality is a matter of reasoning. In his reply, Greene criticizes Mikhail's account of the old cases and presents new and independent evidence for his claims. Then he asks Timmons why deontologists should trust emotions that are fickle and contingent.

Our moral judgments can also be illuminated by comparison with moral judgments by abnormal people, including people with psychopathy, acquired sociopathy, and autism. This method is adopted in the chapters 3 through 5.

Kent Kiehl focuses on psychopathy. Psychopaths are often said to lack conscience or morality because of how they act, but, surprisingly, they do not lack intelligence or the ability to articulate verbally appropriate moral judgments about many real-life situations. To understand this baffling syndrome, Kiehl reviews the neuroscience literature and concludes that psychopathy is associated with dysfunction of the paralimbic system. The particular neural regions implicated include the orbital frontal cortex; insula, amygdala, and parahippocampal regions; anterior superior temporal gyrus; and rostral, caudal, and posterior cingulate.

In their comment, Ricardo de Oliveira-Souza, Fátima Azevedo Ignácio, and Jorge Moll report their unpublished work on “true community antisocials” and ask whether this distinct population fits Kiehl’s model, which is based on violent criminals. Jana Schaich Borg then cites her own work as support for her speculation that psychopaths might resemble normals in some moral judgments but not in others.⁴ She also questions the dichotomy between emotion and cognition that underlies Kiehl’s claim that psychopaths display an intact intellect despite emotional dysfunction.⁵ In his reply, Kiehl reveals further complexities both in the distinction between emotion and cognition and in studies of true community antisocials as opposed to criminal psychopaths.

Psychopathy is often conflated with so-called acquired sociopathy, which is due to damage to the ventromedial frontal lobe, but these syndromes differ significantly in some respects that are important to moral judgment. In chapter 4, Jeanette Kennett and Cordelia Fine discuss these differences and argue that psychopaths do not make moral judgments except in an “inverted commas” sense, but at least some acquired sociopaths are able to make third-person hypothetical moral judgments.⁶ Where acquired sociopaths characteristically fail is in applying those abstract moral judgments to their own situations in the first person. Thus, contrary to published arguments by Adina Roskies, cases of acquired sociopathy do not refute moral internalism—which Kennett and Fine take to be the philosophical thesis that other things being equal, any fully rational person who makes an *in situ* moral judgment is motivated to act accordingly.

Roskies responds by clarifying how her prior argument works against the specific form of internalism that was her target and then criticizes Kennett and Fine’s counterargument that evidence from psychopathy actually supports internalism. Michael Smith argues that Kennett and Fine weaken moral internalism too much when they restrict its claims to *in situ* judgments and admit that it holds only “other things being equal,” and then

Smith defends his own qualified moral internalism against Roskies's criticisms. In reply, Kennett and Fine argue that both sides in the debate over internalism depend on assumptions about which mental states count as moral judgments, so internalism cannot be disproved by any empirical discoveries alone.

Another abnormal syndrome that has attracted attention from moral philosophers is autism. Jeanette Kennett argued that moral judgments by high-functioning autistics support a Kantian rationalist view of moral judgment and agency. In contrast, here Victoria McGeer argues that Humean sentimentalists can accommodate and, indeed, provide a better explanation of moral judgments by individuals with autism. McGeer cites self-reports and other data to show that autistics' moral judgments are based, not on reverence for pure reason but instead on a passion for order. McGeer concludes that disinterested concern can be rooted in a concern for the well-being of others, a concern with social structure and position, or a concern with cosmic structure and position. These three spheres of disinterested concern or varieties of moral agency are all present in normal humans without autism, and autistics share at least some of these concerns.

Kennett responds that McGeer's evidence suggests that Humeans and Kantians have been talking past each other, so Hume's descriptive account and Kant's normative account can be reconciled. Heidi Maibom then presses McGeer for a better explanation of cosmic concern and supports her focus on social concern with additional evidence from the Milgram and Stanford prison experiments. Next, Frédérique de Vignemont and Uta Frith argue for a different view of autism, in which autistic moral judgment and agency are distinctive not because of a lack of empathy or insensitivity to others' distress but instead because of a lack of interaction between an egocentric view of other people (in relation to themselves) and an allocentric view of other people (as having separate lives of their own). In her reply, McGeer elaborates her speculations on the moral importance of cosmic structure and position, and she argues that Frith and de Vignemont's alternative view of autism needs to be supplemented with a fuller account of moral violations in order to cover the variety of cases.

Additional lessons about moral judgment can be gleaned from developmental psychology and neuroscience. Chapters 6 and 7 adopt this perspective, addressing childhood and adolescence in turn.

Jerome Kagan's chapter postulates a universal sequence of stages in the early development of morality. Infants first learn that certain behaviors are followed by punishments, but soon children display a reluctance to violate standards even when they have not experienced prior punishment for

violating that standard. Next, children apply the terms “good” and “bad,” followed by feelings of guilt, and then the abstract concepts of “fairness” and the “ideal.” Kagan emphasizes that this developing morality depends on social categories that have lost much of their moral power in contemporary culture. He then explains how, within these universal patterns, individual variations result from heritable temperaments, including different degrees of vulnerability to guilt, which he traces to patterns of activity primarily in the amygdala.

In their comment, Nathan Fox and Melanie Killen discuss the role of culture in moral development and some pros and cons of the lost power of social categories. Paul Whalen then explains why more work needs to be done on connections between the amygdala and the ventromedial prefrontal cortex. Kagan replies by agreeing that there are costs when social categories lose power and that more than the amygdala is involved in morality.

Chapter 7 turns to a later stage in development—adolescence. Abigail Baird describes four basic stages: classical conditioning first reduces or encourages behaviors by pairing them with sensory outcomes; then through operant conditioning children internalize mental schemas that represent behavioral standards; then more complex abstract thought emerges; and finally comes the sense of belonging to a larger society. With regard to the development of morality during adolescence, Baird emphasizes the roles of cognition, self-conscious emotion, and the transition from a parent-centered to a peer-centered social world. She argues that the integration of visceral emotion with social cognition during adolescence, which is enabled by the maturation of the prefrontal cortex, is essential for a fully developed moral reasoning that functions with minimal cognitive effort. In her view of moral development, knowing precedes feeling, and over time visceral feelings of wrongdoing become rapid automatic responses.

Daniel Lapsley concurs with Baird’s emphasis on automaticity, as opposed to Kohlberg’s phenomenalism, but questions several aspects of Baird’s four-stage model. Katrina Sifferd then asks what Baird’s views on adolescent development imply about the origins of pathologies and about juvenile criminal culpability and punishment. Baird replies by developing her views on pathological disruption in the proposed model of moral development and by explaining how horrific moral transgressions can occur in the absence of any discernable pathology.

Richard Joyce closes this three-volume collection with a sober warning not to become overexuberant in drawing philosophical lessons from empirical findings. Jonathan Haidt and Joshua Greene have sometimes

suggested that neuroscience and psychology support emotivism.⁷ Joyce denies this suggestion if emotivism is understood as the philosophical claim that people who make moral judgments are expressing, not beliefs, but only desires, emotions, or preferences. Similarly, Shaun Nichols argues that experiments on psychopaths and on folk views of psychopaths undermine psychological and conceptual versions of moral rationalism.⁸ Joyce criticizes Nichols's argument and concludes that neuroscience cannot undermine conceptual moral rationalism, much less the claim that moral judgments can be rationally justified. In the end, Joyce agrees that empirical research can be relevant to philosophical moral theory, but not as directly as is often supposed.

Shaun Nichols responds by arguing that conceptual rationalists will have more trouble explaining his experimental results than Joyce admits and then by showing how empirical results can undercut the force of some popular attempts to rationally justify moral judgments. Leonard Katz then outlines a way in which the neuroscience of pleasure and pain, together with other scientific theories, might reveal a source of ultimate and objective normatively justifying reasons. Joyce replies that Nichols's experiments might reveal ordinary opinions without illuminating conceptual content, that justificatory rationalists need not appeal to moral intuitions in the way that Nichols assumes, and that Katz's attempted rationalist justification cannot cross the interpersonal divide and show why facts about pain and other hedonic states give *me* any reason not to cause pain to *you*.

These brief summaries cannot come close to doing justice to the subtlety and depth of the exchanges in this volume. Nor can these exchanges finally solve any of the problems raised in these pages. The most that can be reasonably hoped is that the chapters in this volume and its predecessors in this collection will lead philosophers to become interested in relevant empirical research and will lead psychologists and neuroscientists to do more work on issues related to philosophical problems. Our best hope in moral psychology is for philosophers, psychologists, and neuroscientists to work together. These volumes show that collaboration and discussion among these fields can be fruitful and should continue.

Notes

1. Compare Haidt's chapter in volume 2 of this collection.
2. Greene's claims about the role of moral emotions in moral judgment can be usefully compared with Nichols's chapter in volume 2 of this collection.

3. Mikhail's "moral grammar" approach is a version of the linguistic hypothesis defended by Hauser et al. in their chapter in volume 2 of this collection. That chapter includes further criticisms of Greene's model.
4. Schaich Borg also cites a study reported in the chapter by Hauser et al. in volume 2 of this collection.
5. The chapter by Moll et al. in this volume raises related problems for the dichotomy between emotion and cognition.
6. A new study of moral judgments by patients with frontal lobe damage is reported in the chapter by Hauser et al. in volume 2 of this collection.
7. Compare the chapters by Greene in this volume and by Haidt and Björklund in volume 2 of this collection.
8. See Nichols's chapter in volume 2 of this collection. Compare also the debate among Kennett, Fine, Roskies, and Smith in this volume.

Moral Psychology

