PHILOSOPHY & PSYCHOANALYSIS SERIES. Series Editor: Jon Mills

ETHICS and ATTACHMENT

HOW WE MAKE MORAL JUDGEMENTS

ANER GOVRIN

ROUTLEDGE

"*Ethics and Attachment: How we Make Moral Judgments* is brilliant. Every scholar and researcher concerned with the psychology of morality should read this scientifically sound, theoretically innovative, and gracefully written volume. While applying insights from psychoanalysis, attachment theory, and neuroscience into the morality field, it opens up new avenues of research on the contribution of child's attachment experiences to moral judgments later in life. I thoroughly enjoyed this book as both an attachment researcher and a person concerned with moral wrongs and the promotion of a more just and harmonious world."

Mario Mikulincer, Professor of Psychology,
Interdisciplinary Center (IDC) Herzlyia, Israel

"Govrin explores a revolutionary thesis with tight logic and irrefutable evidence, revealing that even our most complex moral judgments have a surprisingly simple beginning—the dyad of caregiver and infant. Through discussions that range from psychology to philosophy, ethics to evolution, Govrin leaves little doubt that he is a true scholar of morality, and that the elegant "Attachment Approach" to morality helps explain much about our judgments of good and evil. A must-read for anyone interested in moral psychology."

Kurt Gray, Associate Professor of Psychology
and Neuroscience, University of North Carolina
at Chapel Hill, USA

"Drawing on sources as diverse as psychoanalysis and neuroscience as well as philosophy and attachment theory, Dr. Govrin discerns in the earliest years of life the basic cognitive and affective structures undergirding the complexity of moral life. He combines a careful analysis of moral conflicts with a sensitive and thorough exploration of how we evaluate them, both consciously and unconsciously. *Ethics and Attachment: How We Make Moral Judgments* is a masterpiece of creativity and scholarly integration that will be required reading for anyone interested in the psychology of moral judgment."

Ronald C. Naso, psychoanalyst and clinical
psychologist in independent practice, Stamford, CT,
USA; current President, American Board and Academy
of Psychoanalysis

Ethics and Attachment

Why are we disgusted when an elderly woman is robbed but sympathize with the actions of a Robin Hood? Why do acts of cruelty against a helpless kitten bother us more than does the trampling of ants?

In *Ethics and Attachment: How We Make Moral Judgments*, psychoanalyst and philosopher Aner Govrin offers the attachment approach to moral judgment, an innovative new model of the process involved in making such moral judgments.

Drawing on clinical findings from psychoanalysis, neuroscience and developmental psychology, the author argues that infants' experience in the first year of life provides them with the basic tools needed to reach complex moral judgments later in life. With reference to Winnicott and Bowlby, the author examines how attachments affect our abilities to apply to make moral decisions.

With its wholly new ideas about moral judgments, *Ethics and Attachment* will be of great interest to ethics and moral philosophy scholars, law students, and psychoanalytic psychotherapists.

Aner Govrin is a clinical psychologist, psychoanalyst and a director of a doctoral program in the Program for Hermeneutics and Cultural Studies at Bar-Ilan University, Israel. He is a member of the Tel Aviv Institute for Contemporary Psychoanalysis (TAICP).

Philosophy & Psychoanalysis Book Series
Jon Mills
Series Editor

Philosophy & Psychoanalysis is dedicated to current developments and cutting-edge research in the philosophical sciences, phenomenology, hermeneutics, existentialism, logic, semiotics, cultural studies, social criticism, and the humanities that engage and enrich psychoanalytic thought through philosophical rigor. With the philosophical turn in psychoanalysis comes a new era of theoretical research that revisits past paradigms while invigorating new approaches to theoretical, historical, contemporary, and applied psychoanalysis. No subject or discipline is immune from psychoanalytic reflection within a philosophical context, including psychology, sociology, anthropology, politics, the arts, religion, science, culture, physics, and the nature of morality. Philosophical approaches to psychoanalysis may stimulate new areas of knowledge that have conceptual and applied value beyond the consulting room reflective of greater society at large. In the spirit of pluralism, *Philosophy & Psychoanalysis* is open to any theoretical school in philosophy and psychoanalysis that offers novel, scholarly, and important insights in the way we come to understand our world.

Titles in this series:

Temporality and Shame: Perspectives from Psychoanalysis and Philosophy
Edited by Ladson Hinton and Hessel Willemsen

Progress in Psychoanalysis: Envisioning the Future of the Profession
Edited by Steven D. Axelrod, Ronald C. Naso and Larry M. Rosenberg

Lacan on Psychosis: From Theory to Praxis
Edited by Jon Mills and David L. Downing

Ethics and Attachment

How We Make Moral Judgments

Aner Govrin

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2019
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Names: Govrin, °Aner, 1966- author.
Title: Ethics and attachment : how we make moral judgments / Aner Govrin.
Description: 1 Edition. | New York : Routledge, 2018.
Identifiers: LCCN 2018026481| ISBN 9781138079779 (hardback) | ISBN 9781138079786 (pbk.) | ISBN 978131511428 6 (Master) | ISBN 9781351627252 (Web) | ISBN 9781351627245 (epub) | ISBN 9781351627238 (mobipocket)
Subjects: LCSH: Decision making. | Psychology—Moral and ethical aspects.
Classification: LCC BF448 .G69 2018 | DDC 153.8/3—dc23
LC record available at https://lccn.loc.gov/2018026481

ISBN: 978-1-138-07977-9 (hbk)
ISBN: 978-1-138-07978-6 (pbk)
ISBN: 978-1-315-11428-6 (ebk)

Typeset in Times New Roman
by Swales & Willis Ltd, Exeter, Devon, UK

To my mother, Nurit Govrin

Contents

Preface x
Acknowledgments xix

PART 1
Conceptual and empirical foundations 1

1 Why we need a new psychology 3
2 Morality and early interactions: main theories 39
3 The moral skills of infants 67

PART 2
The attachment approach to moral judgment 85

4 The building blocks of moral judgment 87
5 Decoding moral situations 112
6 Variance and consistency in moral judgment 130
7 The like-me criterion and turned-off dyads 165
8 The prototype of evil 194

Epilogue 225
References 229
Index 254

Preface

This book calls for a conceptual revolution in our understanding of morality.

It presents a novel theory, the attachment approach to moral judgment, which explains what stands behind most of the moral judgments we reach. Many attempts have been made to explain the intriguing process of moral judgment. This has produced a variety of answers, some clashing, from a range of disciplines. In what follows, I will attempt to show that a less apparent truth underpins the process of moral judgment and that the explanations so far offered resolve only part of the puzzle.

My main thesis is that moral judgments are based on a unique social cognition. The cognition is so simple that one wonders why no one has identified it until now. In fact, my book only puts into words what every human being knows in their heart of hearts from a very early age, an unconscious relational knowledge, like music that has been playing all our lives without our noticing it.

This provides an answer to an array of seemingly very divergent and baffling questions: Why does cruelty towards a helpless kitten bother us much more than trampling ants? Why are we disgusted when an elderly woman is robbed and yet admire a Robin Hood? How can it be that people who share a general moral code are poles apart when it comes to judging the morality of abortion or capital punishment? Moreover, the principle proposed in this book also addresses such theoretical questions as: What is the nature of ethics? Does it rely on reason or emotion? What is evil?

The theory in a nutshell

People perceive all moral situations, however different, using the same set of parameters. These universally shared criteria enable us to easily identify

moral situations and differentiate them from other non-moral events. At their core, moral judgments require us to *judge relations between two parties*. Moreover, moral judgments are not limited to judging one isolated, single component of a moral situation, such as intentionality or the extent of harm caused, but rather require an assessment of an entire relationship that is held up to our prior expectations of how relationships of this type should be handled.

One of the most important factors in judging relations is assessing the asymmetry of power between the parties.

This is why although robbery is in itself deemed to be wrong, we will unequivocally condemn a thief robbing an elderly woman, but maybe less so someone who defrauds his insurance company, and we might even salute a Robin Hood.

The intent is the same, the action is the same and the damage is the same: an unlawful appropriation of someone else's property. And yet these are three different dyads with three different relationships between victim and thief, resulting in three different judgments.

The picture becomes still more complex when different observers evaluate the same dyad. The thief's mother in the first event is likely to reach a different judgment than will the son of the old woman, and both will differ from the third judgment reached by an independent judge. Thus, three judges, each with differing interests and perspectives, will assess the relations between the sides in a different manner.

Nonetheless, I will argue that even though they reach different conclusions, the three judges are engaged in the same mental actions, using the same parameters, evaluating the same data. Moreover, the three judges are able to reach widespread agreement as to the nature of moral judgment, which enables them to communicate and understand one another. For example, they all agree that one must not steal, and certainly not from the poor, the sick, the oppressed and so on.

The meaning of a moral situation is a function of the meaning of its parts and of the dyadic rules by which they are combined. We break the moral situation down into its most basic component parts. Within the dyad, we identify relations between two sides: strong/weak, dependent/independent, helpless/in control.

We have a range of expectations as to how the strong party to a dyad should and should not behave towards the weak side. We perceive moral failure when, as observers, we believe that the conduct of the strong towards

the weak has violated our expectations. This social cognition is universal. We always expect the strong to protect the weak or at least cause them no harm. This expectation stands behind every person's moral judgment in every culture. In the book, I explain in detail the differences between individuals and between societies related to moral judgment. A side identified by one observer as "weak" will be considered "strong" by another. And yet, even though they reach contradictory conclusions, the two observers analyze the moral situation using the same parameters. They are engaged in the same cognitive calculation: detect a dyad, quickly identify the weak and strong parties, and assess whether and to what extent there was a violation of expectation.

However, even though the cognitive calculation is universal, our relations towards each of the sides, the sympathy and hostility we feel towards them, constitute an unstable and variable set of factors that differ from person to person and from culture to culture.

In this book, I argue that our moral intuitions originate in our expectations regarding the position of the weak. We have this knowledge intuitively, it is accessible to us effortlessly and directly, and we feel that it requires no justification or explanation. One of my main arguments is that this intuitive knowledge is based on our earliest experience, from the first year of life, when we were ourselves part of a dyad in which we were the weak side and our life depended on the devotion, care and protection of the stronger side.

This book's main thesis is that in the first year, a powerful early organizing process takes place that eventually enables the infant to abstract what is common to all moral situations. Thus, we tend to intuitively view moral situations as a perceived interlocking dyadic system of childlike and adult-like parties. Let me explain.

From the moment of birth, and for a considerable time thereafter, we are all nurtured and nursed by a caregiver. Strip humans of all their personal life stories and cultural practices, and you are left with one pivotal formative experience that humans have shared for as long as *Homo sapiens* has inhabited our planet. In the early years of life, every child goes through a period of total dependence, which, if the child is to survive and grow, requires relations with a caregiver. We are born into a dyadic situation: two human beings who, at the nonverbal level, influence each other in a moment-by-moment coordination of the rhythms of behavior, emotion and cognition. The two parties are merged with one another though their functions are distinct: only one of the two is dependent on the other in

sustaining life and growth. This long phase of dependency forms a *unique social cognition, the prototypical dyad*: infant/caregiver, dependent/independent; the relationship between one who has all the power and another who lacks any form of power, one who is needy and one who is resourceful, one who is responsible and one who is not, and so on. Moreover, the prototypical dyad determines not only how these two parties differ; it also provides a clear understanding of their preferred interrelations. The perceived independent ought to be committed to the welfare of the perceived dependent and responsible for the dependent's welfare. The dependent is not obliged to do the same.

In certain dyads, the obligation is total, as in the first primary dyad, between caregiver and infant. In other situations, it may be partial, as in a teacher/pupil dyad or a psychotherapist/patient dyad. Sometimes the obligation is limited to a specific responsibility such as the obligation a policeman has towards a citizen. And sometimes the obligation is simply not to harm, or to be very careful, as in the car driver/pedestrian dyad.

From birth, culture will pull this basic structure in diverse directions. Cultures differ from each other in the objects they define as dependent and independent – what is considered moral in one culture is considered transgressive in another. But there is a unity in the cognitive processes that lead people to interpret and judge moral situations, even though the results may widely vary.

The fundamental structure of this social recognition is innate and becomes activated during early childhood, just like other capacities such as language acquisition. The infant has an innate disposition to participate in a dyad in which a caregiver supplies his or her basic needs. When this innate disposition is brought into contact with a realization that resembles it – a real caregiver who is motivated to meet the infant's needs, even in a minimal way, just to maintain life – this social cognition starts to evolve. One of the ways it evolves is by learning to extend its application to situations outside the dyad, to other group members and even strangers.

Lakoff and Johnson (1999) write:

> Brains tend to optimize on the basis of what they already have, to add only what is necessary. Over the course of evolution, newer parts of the brain have built on, taken input from, and used older parts of the brain.
>
> (p. 43)

Paraphrasing Lakoff and Johnson's theory, we can ask if it is really plausible to suggest that if the infant/caregiver system can be put to work in the service of a parent protecting his or her child, the brain would build a new system to duplicate what it could already do in other social relations?

All humans are capable of distinguishing between good and bad. In saying that, I am asserting that every human possesses the necessary knowledge to reach moral judgments. Apart from those with serious mental handicaps, we all have the ability to break down a moral situation into its basic component parts, identify who is the dependent and who is independent, and analyze the relations between them in terms of violations of expectations.

To perform these mental functions, a person does not need additional instructions, practice or physical bodily development. The natural environment in which the child is raised is sufficient. The infant simply needs to be a part of asymmetrical relations in which someone cares for him or her and ensures that he or she is able to survive at the most basic level. Even an infant to a depressed or abusive caregiver has been given enough care to survive.

None of this infers that education and social interaction are of no importance. On the contrary, the basic innate moral skills expressed reflected in learning the dyadic rules only constitute an initial platform through which the developing child is able to understand and bestow meaning to moral education.

The attachment approach to moral judgment rests on the notion of a "moral intuition" that is unconsciously enacted and automatic. However, just because it is automatic does not make it moral. Moreover, moral intuitions can be entirely relative and subjective based on desire and emotional prejudice.

Morality is typically thought of as a developmental achievement, not something innate, hence it requires socialization and self-consciousness, which an infant does not have.

I do not distinguish between moral judgment and the intuitive capacity to discriminate between good and bad because in my view it is difficult to know where the border between them lies. In this book, I argue that the moral thought which uses principles of justice and morality, such as that applied by a judge in a juridical system based on democratic and liberal principles, also uses the fixed universal parameters that serve our intuitions. Such principles put a constraint on the computation result, but they

do not change the computation process. Thus, in this sense, we are not talking about two separate systems as many mistakenly thought.

*

The book is in two parts. The first part, "Conceptual and empirical foundations," outlines the rationale of the new theory.

In Chapter 1, "Why we need a new psychology," I present a critique of moral psychology as it is practiced mostly by the followers of either Kant's or Hume's philosophies. Moral psychology has so far mainly tended to be dedicated to proving that the essence of morality is either "cognitive" or "emotional." But because both emotion and cognition should be considered equally central to moral psychology, any dichotomizing account will necessarily miss their content. In the final section of this chapter, I describe two theories in moral psychology that took the latter course without pitting cognition against emotion – universal moral grammar (UMG) (Hauser, 2006; Mikhail, 2007, 2012) and the theory of dyadic morality (Gray et al., 2012b; Schein & Gray, 2017).

In Chapter 2, "Morality and early interactions: main theories," I survey the modest body of thought, gleaned mostly from disciplines outside of moral psychology, according to which the origin of our moral capacities resides in the early interactions between infant and caregiver.

Chapter 3, "The moral skills of infants," cites extensive evidence according to which an infant possesses a minimal set of skills required for moral judgment already in the first 12 months of life. It describes the many works of research that relate to infants' ability to understand others' intentions, motivations and desires, to prefer people who resemble them, and to reach basic moral decisions. Such studies offer us an unparalleled opportunity of discovering humans' socio-moral anticipations at the very beginning of life.

The second part of the book, "The attachment approach to moral judgment," describes various aspects of the attachment approach to moral judgment.

Chapter 4, "The building blocks of moral judgment," discusses questions such as: What characteristics do different moral situations have in common? How do people recognize moral situations and identify regularities within them? What are these regularities? My assumption is that we deal with moral situations in the same way we deal with other objects or events. We categorize the situation as moral and then judge it according to

which *pre-existing representation* it most closely resembles. I present the formula A → C as representing the building blocks of all moral situations.

Chapter 5, "Decoding moral situations," shows that we represent each of the parties in ways that are comparable to our fundamental representation of children (or infants) and adults. All our efforts are geared to construct the reality of the moral situation in terms of a child/adult dyad. We also evaluate the relationship between the adult-schema and child-schema parties in terms of their relations (→). We possess a schema for the dyadic relation, centered on our knowledge of adult obligations to children.

Chapter 6, "Variance and consistency in moral judgment," explores the fixed as well as the more variant foundations of moral judgment. The ability of supporters and opponents of abortion or of the death penalty to communicate and understand one another stems from the fact that the moral arguments used are subordinate to fixed and uniform parameters. Disagreements (i.e. different judgments) arise from the different weight that the two opposing sides assign to the fixed parameters. The chapter also provides a coherent account of how we construe moral justifications. It shows that any attempt to change a moral judgment through a rational argument must focus on the perception of the childlike and adultlike features. This holds true not just for concrete cases, but also for moral principles. Extensive feedback loops and robust top-down constraints operate in a way that any information about these features influence the entire gestalt.

Chapter 7, "The like-me criterion and turned-off dyads," describes the subjective nature of moral judgment. We encounter many moral failures around us, but they often leave us indifferent and lacking emotional involvement. The chapter describes the conditions in which the moral judgment is charged with a motivational/emotional component. Through people's attitudes to different animals (some fortunate to possess human characteristics) the like-me criterion is described as the decisive criterion in the process of reaching moral judgments. *Turned-off dyads* are construed when an observer recognizes the moral failure without difficulty and yet the suffering of the victim fails to arouse any emotional or motivational component in him or her.

Chapter 8, "The prototype of evil," implements the attachment approach of moral judgment, to the perception of evil. I suggest that the perception of evil consists of four salient features: extreme asymmetry between victim and perpetrator; a specific perceived attitude of the perpetrator towards the victim's vulnerability; the observer's inability to understand the perpetrator's perspective; and insuperable differences between the observer and

perpetrator's judgment following the incident, which shake the observer no less than the event itself. I then show that the perception of evil involves a cognitive bias. The observer is almost always mistaken in his or her attributions of a certain state of mind to the perpetrator. The philosophical and evolutionary significance of this bias is discussed, as well as suggestions for future testing of the prototype model of evil.

The Epilogue suggests how the insights of the attachment approach to moral judgment might help us to enhance dialogue with people who support conflicting moral views.

The fact of my being a psychoanalyst has had an enormous influence on the writing of this book. In their daily work, analysts always put the dependency and vulnerability of the patient, together with the asymmetrical relations between them, at the center of their understanding and technique. Perhaps more than in any other discipline, psychoanalysis has taken a huge interest in the inner world of infants. In the attempt to understand people and human suffering, psychoanalysis attaches great importance to the fact that we all come into this world in an utterly helpless state and completely dependent on a caring parent.

The child/caregiver dyad is perhaps the most under-theorized domain in moral psychology. Though there is a vast body of literature showing the link between patterns of early attachment and moral behavior (Chugh et al., 2014; Mikulincer & Florian, 2000; Robinson et al., 2015), the subject of parenthood in infancy remains on the margins, exiled by moral psychology's longstanding cultural bias against it and ignorance of the subject's importance.

Researchers working in multiple disciplines have been a source of inspiration in the writing of this book (see Chapter 2): John Bowlby (psychoanalysis), Donald Winnicott (psychoanalysis), Carol Gilligan (feminist thought), Anton Dijker (social psychology), Darcia Narvaez (psychology) and Patricia Churchland (philosophy). These theorists were the first to reveal how psychic life in infancy paves the way to our moral capacities. While all these theories are well known and understood in their respective fields, very few of them have played a part in, or influenced, moral psychology.

In this book, I offer critiques of other thinkers in the field of moral psychology. Chapter 2 in particular makes me uneasy because there is always the possibility that I have not properly understood their ideas. In my defense, I can only say that I have done my best to offer a fair summary of a whole family of ideas that I found to be too restrictive.

This book aims to show how behind a vast variety of opinions, similar cognitive processes are at play. Always, moral judgment involves moral calculations, and these in their turn – no matter how complex the situation at hand – hark back to the same basic parameters experienced by each and every human being during the first year of life.

The attachment approach to moral judgment is an overarching theory that explains distinct domains of morality that seems unrelated. To expose the laws and working principles that appear consistently across organizational levels from the largest to the smallest, and to discover regularity behind phenomena that appear to be disparate and dissimilar, is the grand objective of science.

Of course, we must continue to honor the uniqueness of every field, since the formulation of principles at the simplest level is not always sufficient if we are to understand the more complex moral phenomena. I hope that researchers in the fields of psychoanalysis, moral psychology, moral philosophy, anthropology, neuroscience and sociology will continue to expand the meeting ground between cognition, emotion and attachment. Conceivably, this encounter will lead to change in each of the separate fields of research, will shed new light on them all, and broaden the domain in which their findings can be implemented in a way that goes beyond their original objectives.

Acknowledgments

I gratefully acknowledge the advice and assistance of Oren Harman, Mordechai Nisan, Golan Shahar and David Bar-Gal, who read sections of the book and offered their wise comments.

Kurt Gray, whose ideas have been inspiring in the writing of this book, invited me to write a chapter on moral psychology in a book he jointly edited with Jesse Graham titled *Atlas of Moral Psychology* (Govrin, 2018). That invitation encouraged and stimulated me to write this book. Prof. Ed Greenstein and Prof. Elda Weizmann from Bar-Ilan University have been friends and a tremendous source of support throughout.

Part of the thesis presented here was published by *Frontiers in Psychology* (Govrin, 2014a, 2014b, 2018), and I thank *Frontiers* reviewers and editor Prof. Rafael Malach from the Weizmann Institute of Science, Israel, Rina Lazar from the Sackler Faculty of Medicine, Tel Aviv University, Israel, and David Smith from Robert Gordon University, United Kingdom, for their helpful comments and suggestions.

This book would not have been published without the support and encouragement of Jon Mills, editor of the Routledge series. For me, Jon has been a source of inspiration, guiding me in making rich and fascinating links between psychoanalysis and other fields of study.

A special thank you to Mirjam Hadar, who helped me with the editing of the book, offering me her rich experience and wisdom. My thanks to Dan Gillon, my long-standing translator, who worked on the book with exceptional devotion, and Noga Galor-Ariel, my research assistant, who worked tirelessly and professionally in preparing the book for publication.

I want to thank Kate Hawes of Routledge, who provided a fast and generous response to every question. It is a great privilege for an author to work with a publisher who believes in the books they publish. My thanks

also to Charles Bath of Routledge, who accompanied the text at every stage of production with great care and devotion.

My good friend Baruch Many showed great interest in my ideas and supported me in moments of crisis.

I thank my wife, Osnat, and my children, Or and Yonatan, who are the joy of my life. Over many years, I have had countless conversations with Osnat about the internal structure of moral situations, which clarified many aspects of my theory and helped me to improve it.

I owe my mother, Nurit Govrin, to whom this book is dedicated, a debt greater than I am able to describe here. Her unswerving support of my writing and the way in which she was always there for me enabled me to be creative and write.

I was able to discuss the ideas in this book with my beloved father, Shlomo Govrin, before his death. He was a source of inspiration, support and encouragement throughout my life, in addition to the many insights he offered me in my writing.

Part I

Conceptual and empirical foundations

Chapter 1

Why we need a new psychology

The last decade has witnessed significant progress in the field of moral psychology. Recent research in moral psychology has produced strong evidence to suggest that moral judgment is intuitive, and is accomplished by a rapid, automatic and unconscious psychological process (Damasio, 1994; Greene & Haidt, 2002; Hauser, 2006; Mikhail, 2000; Shweder & Haidt, 1994). This line of research challenged the long-dominant cognitive development paradigm conceived by Kohlberg (Kohlberg, 1969; Piaget, 1932/1965; Turiel, 1983, 2006), according to which moral judgment is the product of conscious, effortful reasoning.

Nonetheless, most theories in moral psychology remain general and undefined, leaving key questions unanswered.

There is, for example, considerable disagreement and confusion as to what moral intuitions are and how they work: What exactly are the underlying cognitive processes of these judgments that "operate quickly effortlessly and automatically, such that the outcome but not the process is accessible to consciousness?" (Haidt, 2001, p. 818). How are moral situations represented in our minds? What cognitive processes intuitively glue together different moral situations to one category?

Most theories did not usually concern themselves with these questions. It is my strongly held view that, with very few exceptions, up until recently all the parties currently engaged in debating the central issues in moral psychology are captive to the eighteenth-century controversy between rationalists and sentimentalists. Indeed, many theories of moral psychology expressly associate themselves with one of these two major moral philosophies.

Astonishingly, moral psychology is still overshadowed by a dispute that dominated this field of study 200 years ago.

Philosophical views about morality understandably affect those whose main occupation it is to develop a theory of moral psychology. Any theorist's explanation of morality is of necessity marked by his or her views about how morality works, and these views were mostly defined by the long tradition of moral philosophy.

Theories in the relatively new, almost fledgling field of moral psychology therefore unsurprisingly refer to key issues raised by philosophers.

One problem, however, that has been the outcome of the relative wealth and depth of the philosophical discourse on the subject compared to its psychological counterpart is that the problems of moral psychology have remained static over time, and so, of course, did the solutions.

In many ways, the main questions have not changed since the publication in 1751 of Hume's *An Enquiry Concerning the Principles of Morals* (1751/1998), in which he noted that: "There has been a controversy started of late much better worth examination, concerning the general foundation of morals; whether they be derived from reason, or from sentiment" (EPM 1, p. 3). This debate, according to Hume, questions whether we can "attain the knowledge of" moral distinctions "by a chain of argument and induction" or whether we must experience "an immediate feeling and finer internal sense" (EPM 1, p. 3).

Kant, by contrast, posited the universal nature of moral judgments. What one ought to do now, or what one thinks to be right or wrong now, is what one should always do, or consider right or wrong. Moral judgments are applicable beyond the specific case, in a way in which emotions are not. The specific judgments are a product of invariant moral principles. And these are derived by and from reason. Only humans are rational and can have principles, though some animals can have feelings and even express them.

Why have most philosophers and psychologists assumed that it is important and appropriate to stress the reason/emotion binary when addressing moral judgments? The answer is that this binary reflects a very profound trait of morality – its dual nature.

Moral judgments depend on cultural and social contexts. They vary over time and from one individual to another. They bear a direct connection to the emotions. Considered from this perspective, moral judgments may appear subjective, arbitrary and very similar to aesthetic judgments or even to our preferences in taste in matters of food. When we compare different cultures or different historical eras, we witness the enormous

variability in terms of morality. This easily leads to the conclusion that there is no regularity or consistency in the domain of moral judgment.

We may, however, choose to describe moral judgments in an essentially different manner. For they don't *feel* at all the same as aesthetic judgments, say of a painting or of a certain dish. In the case of moral judgment, people are not merely affected by their feelings or by the common values of his or her social setting. Often people feel they are unequivocally right, and that their stance is non-negotiable. People believe they are acting on the strength of an objective principle, which holds over and beyond the particular case at hand, to any other situation or question of a similar kind.

No moral philosophy or psychology is complete without a thorough discussion of this duality. The dichotomy between reason and emotion is a reflection of this problem.

However, sticking to it left too many problems unanswered.

For example, throughout this chapter, I will show why premises of followers of each approach – such as "Moral judgments are strictly cognitive and uninfluenced by emotions" or "We cannot form moral judgments without moral emotions" – are inconsistent with empirical findings.

This line of questioning suggests that the two views, when taken in isolation, leave something out – just as rationalists left out emotion, so sentimentalists fail to include cognition.

As I will show, the dichotomous – and often antagonistic – view of emotion and cognition is under increasing challenge. Moral judgments cannot simply be slotted into compartments of the brain related to either cognition or emotion. In the process of reaching moral judgments, affect is neither independent of nor prior to cognition. In the making of an ethical judgment, it is likely that perception and cognition are directly and powerfully influenced by information with affective or motivational content.

The cognition versus emotion debate led most researchers to neglect the patterns and regularities of moral situations and to overlook such important issues as, for instance, how moral situations are represented in our minds and what kind of information the brain encodes and computes when making right/wrong judgments.

Following a brief survey of the controversy between those who support the primacy of cognition in decision-making and those who promote the centrality of emotions, I shall discuss a number of key theories in moral psychology representative of either side of this divide.

The present aim, then, is to illustrate that neither one of these approaches is sufficient, and to show why basing a moral psychology on the underlying binary offers us answers and results that are too restricted.

This chapter does not merely review the main theories in the field, but also shows how the model I propose responds to various theoretical lacunas in this area of study. The paradigm suggested in this book joins a modest tradition of theories that refuse to think of moral judgment as emerging from a binary split between emotion and cognition, and instead search for deep structures that apply specifically to morality.

In the last section of this chapter, I will describe two theories in moral psychology that took the latter course without pitting cognition against emotion – universal moral grammar (UMG) (Hauser, 2006; Mikhail, 2007, 2012) and the theory of dyadic morality (Gray et al., 2012b; Schein & Gray, 2017).

Emotion versus cognition

The emotion versus cognition controversy is at the heart of an old and sometimes passionate debate. It is a polemic that brings to mind the ancient argument as to whether emotion *precedes* and is *independent of* cognition, or whether all forms of emotional expression rely on, or are embedded in, prior processes of thought. In the history of Western thought, reason and affect were viewed as two deeply intertwined yet separate domains. They were thought to be in competition, at times verging on outright antagonism.

But new findings on the cognitive unconscious that takes part in almost every cognitive process, including those at work in complex decision-making, have subverted the common distinction between emotion and cognition. An early version of such a theory is to be found in Leventhal and Scherer (1987). In their view, affective responses, on the one hand, and involuntary, instinctive, non-conscious responses, on the other, should not be differentiated. Leventhal and Scherer regard emotion as the product of an amalgam of components. They argue that the link between emotion and cognition needs to be seen within a theoretical framework that includes a multiplicity of processing components, and which views affect as the product of a combination of these elements. They believe that we no longer need to preoccupy ourselves with defining emotion and cognition separately. Instead of asking "What is emotion?" and "What is cognition?" we

need to identify the part played by certain specific processing components involved in a particular emotional experience.

This early model demonstrates that the split occurs, rather than between emotion and cognition, between emotions *in conjunction* with automatic, non-conscious, pre-reflective processes, on the one hand, and cognitions that stem from controlled processes, on the other.

Research in cognitive science over the past 20 years or so has led to a new view of the unconscious as being capable of functioning in areas previously considered the exclusive domain of conscious thought. These areas include such functions as complex information processing, behavioral patterns, the pursuit of objectives and self-regulation. In his groundbreaking paper "The Cognitive Unconscious" (1987), John Kihlstrom described how early models of human cognition based on the modern high-speed computer had viewed the unconscious as a reservoir of pre-attentive perceptions and dormant traces of memory. According to Kihlstrom, these models erred in assuming that complex mental processes required conscious awareness. Unlike the psychoanalytic unconscious, the cognitive unconscious includes no inherent drives seeking satisfaction. According to Kihlstrom, when compared to the intensity and illogicality of psychoanalytic drives and conflicts, the cognitive unconscious is somewhat unimpassioned, seemingly rational and lacking motivation. Kihlstrom (1987) thought that "conscious awareness . . . is not necessary for complex psychological functioning" (p. 1450). His model of the cognitive unconscious is mostly concerned with affect, motivation, and even control and metacognition. The cognitive unconscious includes the causal nature of phenomenal experience, of intentionality and free will, and the attribution of these qualities to others as well. As Uleman (2005) put it, "the list of psychological processes carried out in the new unconscious is so extensive that it raises two questions: What, if anything cannot be done without awareness? What is consciousness for?" (p. 6).

The interaction between affect and cognition

The question, then, is about the nature of the interaction in the mind between automatic, rapid, unconscious processes, and slower controlled cognitive processes. Are they antagonistic? Pulling in different directions? Or do they, in fact, complement one another?

This dichotomized view of the brain is currently under challenge by an approach that places the emphasis on the deep integration between regions of the brain, and suggests a holistic model in which none of the brain's regions is categorized as uniquely cognitive or affective (Feldman Barrett et al., 2007; Ochsner & Gross, 2005; Pessoa, 2008).

Recent research such as Goldberg et al. (2014) takes what has been termed a "constructionist" approach, with the central idea that knowledge and meaning are "constructed" from experience. The theory posits that numerous functional networks across the brain are reciprocally activated, giving rise to a gamut of emotional states (Barrett & Bar, 2009; Lindquist et al., 2012). The suggestion that there is a connection between a broad range of cortical functions and emotion is not new. This link is reflected in theories that contend that when emotionally processing an event, an individual's responses are akin to "preparation for action." Indeed, the word "motivation" derives from the Latin *emovere* (literally: to move out). The theory links emotional processing to "action preparation." The claim has been supported by a number of behavioral and physiological studies (Cacioppo et al., 1986; Davidson et al., 2000; Frijda, 1986; Lang et al., 1990; Lang et al., 1997).

Emotional processing therefore does not only involve isolated and separate emotions, but consists of a kind of plan of action – a plan in which both emotions and cognition feature. Luiz Pessoa (2013, 2015) is one of the most prominent advocates of the approach, asserting that cognitive and emotional processes are not separate and dichotomous.

The core argument of his book *The Cognitive-Emotional Brain* (2013) is that cognition and emotion are integrated and highly interdependent. In Pessoa's view, the classical model of a dichotomized brain has been unhelpful to neuro-scientific research. According to Pessoa, a new way of thinking is needed that, instead of dichotomizing brain functions, views them as complementary, each defining the other, and most importantly as not mutually exclusive. Understanding it as a network implies that the mind-brain is not an organ whose functioning can be split between emotion (or motivation) and cognition. Which is to say that emotion and cognition should be viewed as neurally controlled less by properties inherent to particular regions of the brain and more by communications between and among numerous areas within it. From this perspective, affect and cognition are operationally integrated systems that constantly impact on one another.