Is the science of moral cognition usefully modeled on aspects of Universal Grammar? Are human beings born with an innate “moral grammar” that causes them to analyze human action in terms of its moral structure, with just as little awareness as they analyze human speech in terms of its grammatical structure? Questions like these have been at the forefront of moral psychology ever since John Mikhail revived them in his influential work on the linguistic analogy and its implications for jurisprudence and moral theory. In this seminal book, Mikhail offers a careful and sustained analysis of the moral grammar hypothesis, showing how some of John Rawls’ original ideas about the linguistic analogy, together with famous thought experiments like the trolley problem, can be used to improve our understanding of moral and legal judgment. The book will be of interest to philosophers, cognitive scientists, legal scholars, and other researchers in the interdisciplinary field of moral psychology.

John Mikhail is Professor of Law and Philosophy at Georgetown University.
Elements of Moral Cognition

Rawls’ Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment

JOHN MIKHAIL

Georgetown University
To Sarah, Hannah, and Andrew; and to the memory of my parents,
Ramzy and Maryse Mikhail
Homo homini lupus;* who has the courage to dispute it in the face of all the evidence in his own life and in history?

– Sigmund Freud, Civilization and Its Discontents

For when the Gentiles, which have not the law, do by nature the things contained in the law, these, having not the law, are a law unto themselves.... They show that what the law requires is written on their hearts, their conscience bearing witness....

– St. Paul, Letter to the Romans

What we have been saying would have a degree of validity even if we should concede that which cannot be conceded without the utmost wickedness, that there is no God, or that the affairs of men are of no concern to Him.

– Hugo Grotius, On the Law of War and Peace

* Man is to man a wolf.
## Contents

*List of Tables and Figures*  
*Preface*

**PART ONE: THEORY**

1. The Question Presented  
   2.1 Nine Comparisons between Linguistics and Moral Theory  
     2.1.1 The Main Questions  
     2.1.2 The General Answers  
     2.1.3 The Fundamental Arguments  
     2.1.4 The Competence–Performance Distinction  
     2.1.5 The Distinction between Operative and Express Principles  
     2.1.6 Levels of Empirical Adequacy  
     2.1.7 Two Additional Questions  
     2.1.8 Commonsense and Technical Concepts of Language and Morality  
     2.1.9 Theoretical Goals  
   2.2 Preliminary Clarifications about Rawls’ Linguistic Analogy  
   2.3 Outline of Remaining Chapters  
3. The Basic Elements of Rawls’ Linguistic Analogy  
   3.1 Eight Features of Rawls’ Conception of Moral Theory  
     3.1.1 The Argument for Moral Grammar  
     3.1.2 The Problem of Descriptive Adequacy  
     3.1.3 The Distinction between Descriptive and Observational Adequacy
viii

3.1.4 The Distinction between Operative and Express Principles 50
3.1.5 The Distinction between Descriptive and Explanatory Adequacy 51
3.1.6 The Competence–Performance Distinction 51
3.1.7 The Theory-Dependence of the Competence–Performance Distinction 55
3.1.8 The Importance of Idealization 56
3.2 Further Clarifications about Terminology 57
3.3 Moral Theory as a Theory of I-Morality 63
3.4 Some Further Remarks about the Linguistic Analogy 67
3.5 The Contrast with Particularism 71

PART TWO: EMPIRICAL ADEQUACY

4 The Problem of Descriptive Adequacy 77
  4.1 The Trolley Problems 78
  4.2 The Properties of Moral Judgment 82
  4.3 Framing the Problem of Descriptive Adequacy 85
  4.4 Locating the Problem within the Framework of Cognitive Science 87
    4.4.1 Perceptual and Acquisition Models 88
    4.4.2 The Hypothetico-Deductive Method 91
  4.5 Objections and Replies 94

5 The Moral Grammar Hypothesis 101
  5.1 Some Initial Evidence 104
  5.2 Simplifying the Problem of Descriptive Adequacy 106
    5.2.1 Twelve New Trolley Problems 106
    5.2.2 Twelve Considered Judgments 110
  5.3 The Poverty of the Perceptual Stimulus 111
    5.3.1 Labeling the Stimulus 111
    5.3.2 Expanded Perceptual Model 111
  5.4 Outline of a Solution 117
    5.4.1 Deontic Rules 117
    5.4.2 Structural Descriptions 118
    5.4.3 Conversion Rules 120
  5.5 Intuitive Legal Appraisal 121

6 Moral Grammar and Intuitive Jurisprudence: A Formal Model 123
  6.1 Three Simplifying Assumptions 124
  6.2 Structural Descriptions I: Acts, Circumstances, and Intentions 125
    6.2.1 Acts and Circumstances 125
    6.2.2 K-Generation and I-Generation 130
  6.3 Deontic Rules 132
Contents

6.3.1 The Principle of Natural Liberty 132
6.3.2 The Prohibition of Battery and Homicide 133
6.3.3 The Self-Preservation Principle 136
6.3.4 The Moral Calculus of Risk 137
6.3.5 The Rescue Principle 144
6.3.6 The Principle of Double Effect 148
6.4 Structural Descriptions II: A Periodic Table of Moral Elements 153
6.5 Conversion Rules 162
6.5.1 Temporal Structure 171
6.5.2 Causal Structure 172
6.5.3 Moral Structure 172
6.5.4 Intentional Structure 172
6.5.5 Deontic Structure 173
6.6 A Brief Note on Enlightenment Rationalism 174
6.7 Further Clarifications about Act Trees 175
6.8 Concluding Remarks 178

PART THREE: OBJECTIONS AND REPLIES

7 R. M. Hare, Peter Singer, and the Distinction between Empirical and Normative Adequacy 183
7.1 Hare’s and Singer’s Criticisms of Rawls’ Linguistic Analogy 183
7.2 Empirical and Normative Adequacy in Grounds 191
7.3 Empirical and Normative Adequacy in Outline 195
7.4 Empirical and Normative Adequacy in A Theory of Justice: Reflective Equilibrium 197
7.4.1 The Main Contractual Argument of A Theory of Justice 198
7.4.2 The Concept of Reflective Equilibrium 202
7.5 Empirical and Normative Adequacy in Independence 213
7.6 Some Clarifications about Metaethics 217
7.7 Objections and Replies 221
7.8 Summary 227

8 Thomas Nagel and the Competence–Performance Distinction 228
8.1 Nagel’s Criticisms of Rawls’ Linguistic Analogy 228
8.2 Analysis of Nagel’s Arguments 231
8.2.1 The Intuitions of Native Speakers Are Decisive in Linguistics 232
8.2.2 The Moral Intuitions of Ordinary Persons Are Not Decisive in Moral Theory 236
8.2.3 Whatever Native Speakers Agree on Is English 238
8.2.4 Whatever Ordinary Individuals Agree in Condemning Is Not Necessarily Wrong 240
## Contents

8.2.5 The Plausibility of an Ethical Theory Can Change Our Moral Intuitions, but the Plausibility of a Linguistic Theory Cannot Change Our Linguistic Intuitions 248

8.2.6 In Linguistics, Unlike Ethics, the Final Test of a Theory Is Its Ability to Explain the Data 257

8.3 Objections and Replies 258

8.4 Summary 264

9 Ronald Dworkin and the Distinction between I-Morality and E-Morality 266

9.1 Dworkin's Analysis of Rawls' Linguistic Analogy 268

9.1.1 The Natural Model versus the Constructive Model 269

9.1.2 The Natural Model and Reflective Equilibrium Are Incompatible 271

9.2 Problems with Dworkin's Analysis 274

9.3 Dworkin's Misinterpretations of Rawls 276

9.3.1 Subject Matter 277

9.3.2 Goal 280

9.3.3 Evidence 282

9.3.4 Method 287

9.4 Objections and Replies 291

9.5 Brief Remarks about Moral Grammar and Human Rights 295

9.6 Summary 303

PART FOUR: CONCLUSION

10 Toward a Universal Moral Grammar 307

Appendix: Six Trolley Problem Experiments 319

Bibliography 361

Index 393
Tables and Figures

**TABLES**

1.1 Some Modern Authors Who Draw a Linguistic Analogy (1625–2000)  
2.1 Seven Main Problems in the Theory of Moral Cognition  
3.1 The Basic Elements of Rawls’ Linguistic Analogy  
3.2 The Heart of Rawls’ Linguistic Analogy  
4.1 Considered Judgments Elicited by Six Trolley Problems  
5.1 Twelve Trolley Problems  
5.2 Twelve Considered Judgments  
5.3 The Poverty of the Perceptual Stimulus  
6.1 Purposely and Knowingly Harmful Acts and Omissions  
6.2 A Periodic Table of Moral Elements (Version 1)  
6.3 A Periodic Table of Moral Elements (Version 2)  
6.4 Manipulating Structural Features: Fourteen New Trolley Problems  
9.1 Dworkin’s “Natural” and “Constructive” Interpretations of Rawls’ Conception of Moral Theory  
A.1 Justifications for the Bystander and Footbridge Problems by 10 Subjects in Experiment 2 (Within-Subject Design)  
A.2 Explaining Six Trolley Problems as a Function of the Principle of Double Effect  
A.3 Explaining Six Trolley Problems as a Function of the Pauline Principle

**FIGURES**

4.1 Simple Perceptual and Acquisition Models for Language and Morality  
5.1 Deontic Concepts and Deontic Logic: (a) Deontic Modality in Natural Language; (b) Square of Opposition and Equipollence
## List of Tables and Figures

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>Classifying the Stimulus: (a) Scheme; (b) Application to the Bystander Problem</td>
<td>112</td>
</tr>
<tr>
<td>5.3</td>
<td>Two Inadequate Appraisal Theories: (a) Unanalyzed Link in Haidt's (2001) Model of Moral Judgment; (b) Inadequacy of Greene's (2001, 2002) Personal–Impersonal Distinction</td>
<td>113</td>
</tr>
<tr>
<td>5.4</td>
<td>Expanded Perceptual Model for Moral Judgment</td>
<td>114</td>
</tr>
<tr>
<td>5.5</td>
<td>Structural Descriptions of Action Plans in Footbridge and Bystander Problems</td>
<td>118</td>
</tr>
<tr>
<td>5.6</td>
<td>Alternative Structural Descriptions of Action Plans in Footbridge and Bystander Problems</td>
<td>119</td>
</tr>
<tr>
<td>6.1</td>
<td>Circumstances Alter Cases: (a) Variance in Six Trolley Problem Experiments; (b) Six Structural Descriptions; (c) An Illustration of Mill's (1843) “Mental Chemistry”</td>
<td>127</td>
</tr>
<tr>
<td>6.2</td>
<td>Computing Structural Descriptions</td>
<td>168</td>
</tr>
<tr>
<td>6.3</td>
<td>Moral Geometry: Varying Structural Descriptions of Four Identical Sets of Act-Type Descriptions</td>
<td>174</td>
</tr>
<tr>
<td>7.1</td>
<td>Schematic Diagram of Rawls' Definition of Reflective Equilibrium in Section 4 of <em>A Theory of Justice</em></td>
<td>206</td>
</tr>
<tr>
<td>7.2</td>
<td>Schematic Diagram of Rawls’ Account of the Problem of Normative Adequacy in <em>A Theory of Justice</em></td>
<td>209</td>
</tr>
<tr>
<td>8.1</td>
<td>Chomsky's (1964) Perceptual Model</td>
<td>253</td>
</tr>
<tr>
<td>8.2</td>
<td>Three Formats for Representing “Ian will throw the man onto the track”: (a) Tree Diagram; (b) Rewrite Rules; (c) Labeled Brackets</td>
<td>254</td>
</tr>
<tr>
<td>9.1</td>
<td>Hohfeld’s (1913, 1917) Fundamental Legal Conceptions</td>
<td>301</td>
</tr>
<tr>
<td>A.1</td>
<td>Moral Judgments of Two Act Types in Experiment 1 (Purposeful Battery vs. Knowing Battery)</td>
<td>322</td>
</tr>
<tr>
<td>A.2</td>
<td>Moral Judgments of Two Act Types in Experiment 1 by Gender (Purposeful Battery vs. Knowing Battery)</td>
<td>322</td>
</tr>
<tr>
<td>A.3</td>
<td>Moral Judgments of Two Act Types in Experiment 2 (Purposeful Battery vs. Knowing Battery)</td>
<td>327</td>
</tr>
<tr>
<td>A.4</td>
<td>Moral Judgments of Two Act Types in Experiment 2 by Gender (Purposeful Battery vs. Knowing Battery)</td>
<td>327</td>
</tr>
<tr>
<td>A.5</td>
<td>Moral Judgments of Two Act Types in Experiment 2 (Good Effects vs. Bad Effects)</td>
<td>328</td>
</tr>
<tr>
<td>A.6</td>
<td>Moral Judgments of Two Act Types in Experiment 2 by Gender (Good Effects vs. Bad Effects)</td>
<td>328</td>
</tr>
<tr>
<td>A.7</td>
<td>Moral Judgments of Two Act Types in Experiment 2 (Purposeful Battery vs. Implied Consent)</td>
<td>329</td>
</tr>
<tr>
<td>A.8</td>
<td>Moral Judgments of Two Act Types in Experiment 2 by Gender (Purposeful Battery vs. Implied Consent)</td>
<td>329</td>
</tr>
<tr>
<td>A.9</td>
<td>Moral Judgments of Two Act Types in Experiment 3 (Purposeful Battery vs. Knowing Battery)</td>
<td>334</td>
</tr>
</tbody>
</table>
List of Tables and Figures

A.10 Moral Judgments of Two Act Types in Experiment 3
(Good Effects vs. Bad Effects) 335
A.11 Moral Judgments of Two Act Types in Experiment 3
(Purposeful Battery vs. Implied Consent) 335
A.12 Moral Judgments of Two Act Types in Experiment 4
(Purposeful Battery vs. Knowing Battery) 339
A.13 Moral Judgments of Two Act Types in Experiment 4
by Gender (Purposeful Battery vs. Knowing Battery) 340
A.14 Moral Judgments of Two Act Types in Experiment 5
(Purposeful Homicide vs. Knowing Homicide) 344
A.15 Moral Judgments of Two Act Types in Experiment 5
(Better Alternative vs. Knowing Homicide) 344
A.16 Moral Judgments of Two Act Types in Experiment 5
(Disproportional Death vs. Knowing Homicide) 345
A.17 Moral Judgments of Two Act Types in Experiment 6
(Purposeful Battery vs. Knowing Battery) 348
At first glance, Freud and St. Paul offer two competing accounts of human nature. On Freud's view, human beings are essentially predatory toward one another. On St. Paul's more optimistic conception, principles of justice and fairness spring from their nature as social creatures. The apparent conflict between these two familiar accounts can be reconciled, however, by means of a competence–performance distinction. Freud describes how humans often do in fact behave toward one another. Writing a decade before the most vicious mass murder machine in history was unleashed on its defenseless victims, he correctly observes that under some circumstances "men are not gentle, friendly creatures wishing for love, who simply defend themselves if they are attacked … [but] savage beasts to whom the thought of sparing their own kind is alien." St. Paul focuses on different behaviors, however, and he also seems correct to infer from them that certain basic norms of conduct are engraved in the mind as a kind of innate instinct. This influential idea, although promoted by the Stoics and embraced by many Jewish and Christian writers, is neither exclusively Western nor inherently religious. Substantially the same notion can be found in the Hindu concept of dharma, the Confucian concept of li, and the writings of the Mu'tazilites and other Islamic rationalists, for example. Likewise, the hypothesis of an innate moral faculty is what supplies the foundation of the jus gentium or law of nations in its traditional, secular formulation. In 1625, Hugo Grotius set the tone for the modern scientific analysis of these subjects and thereby heralded the emancipation of ethics and jurisprudence from theology with his famous etiamsi daremus remark ("even if we should concede") in the Prolegomena to the Law of War and Peace, asserting that a natural moral law would exist even if there were no God, or human affairs were of no concern to him.

This book seeks to revive and develop aspects of the humanistic enterprise pioneered by Grotius and other Enlightenment philosophers by describing and explaining elements of moral cognition within a modern cognitive science framework. The linchpin is the analogy between rules of justice and rules of
grammar, and the gnomon or measuring device is the trolley problem and other artfully designed thought experiments. Just as repeated observations of the gnomon’s shadow gave birth to the science of astronomy by enabling ancient astronomers to compile and organize vast amounts of information about the daily and annual variation of the sun (Kuhn 1957), so too can repeated observations of the moral capacities of human nature as reflected in a variety of common moral judgments provide a secure foundation for moral theory. This, at any rate, is the guiding assumption of the research program described in these pages.

This book began as my Ph.D. dissertation, “Rawls’ Linguistic Analogy: A Study of the ‘Generative Grammar’ Model of Moral Theory Described by John Rawls in A Theory of Justice,” which was submitted to the Department of Philosophy at Cornell University in 2000. Most of the original research was done at Harvard University and the Massachusetts Institute of Technology from 1995 to 1999, beginning with a paper on moral competence I wrote for Noam Chomsky in the spring of 1995 and a series of illuminating conversations about moral theory I held with John Rawls later that summer, before the unfortunate event of his first stroke. The book also draws from the initial trolley problem studies I conducted at MIT from 1995 to 1999, when I was a visiting researcher in Elizabeth Spelke’s Infant Cognition Lab. Many of the book’s main ideas are therefore nearly 15 years old, although they have begun only recently to receive widespread attention, due in part to their popularization by other writers and to the growing use of trolley problems and similar cases to investigate the nature of human moral intuitions.

When I was in graduate school, moral psychology was not a thriving academic discipline, and moral philosophy was for the most part resolutely anti-empirical. As a result, I often found it difficult to persuade philosophers that combining some of Rawls’ and Chomsky’s theoretical insights with actual experiments on people’s moral intuitions was a worthwhile dissertation topic. That would be psychology, not philosophy, is the essence of what I was told. Fortunately, things have now changed, and a new generation of philosophers that seeks to “introduce the experimental method of reasoning into moral subjects” (Hume) and thereby make contributions to cognitive science is receiving greater encouragement. Meanwhile, many psychologists now pursue research on moral cognition along the lines outlined in this book and related publications.

Legal theory has also undergone a transformation in the past decade. Many legal scholars now pay closer attention to the cognitive and brain sciences, and jurisprudence is gradually returning to its naturalistic roots as the attempt to systematize and explain the human sense of justice with the aid of a technical legal vocabulary (as both Adam Smith and Thomas Reid credited Grotius with doing; see The Theory of Moral Sentiments, VII.iv.37, and Essays on the Active Powers of the Human Mind, III). Likewise, cognitive scientists are increasingly drawing on the highly refined concepts of moral and legal theory in the design and analysis of their experiments. In light of all these developments, it
is gratifying to observe that one of the primary motivations of the linguistic analogy – to help promote greater collaborative and interdisciplinary research on human moral capabilities – has already begun to be realized.

It seems unnecessary to observe that much has been learned in the past decade about the topics discussed in this book. In editing and revising the original manuscript for publication, I have not attempted to incorporate what has been accomplished during the intervening period in any systematic fashion. Rather, with a few exceptions, I have sought to preserve and extend the basic structure and argument of Rawls’ Linguistic Analogy as much as possible, despite certain inherent limitations of this procedure. If I were to begin again today from scratch, I would probably write a very different book, one geared less toward fitting my own ideas about moral theory into the format of Rawls’ brief remarks on the linguistic analogy in A Theory of Justice and more toward the independent development of a naturalistic moral psychology within broad scientific parameters. On balance, however, I continue to believe that there are significant advantages to the synthetic approach adopted here, which begins with certain well-known ideas of two seminal thinkers and seeks to integrate and build upon them.

In addition to Rawls and Chomsky, the framework within which this study unfolds draws heavily on the work of Alan Donagan, Philippa Foot, Alvin Goldman, Elizabeth Spelke, Judith Jarvis Thomson, and other highly original researchers, and some familiarity with the fields of linguistics and cognitive science, on the one hand, and moral theory and jurisprudence, on the other, is presupposed. Because I hope to reach a wide and diverse audience, I have nonetheless tried whenever possible to write in plain English and avoid unnecessary academic jargon. However unrealistic, my role models in this regard have been such lucid writers as Descartes, Hume, Mill, and Russell, and as Einstein would have it, my aim throughout the book, however imperfectly realized, has been to render things as simple as possible, but no simpler. At the same time, because of their exceptional value in the clarification and development of scientific ideas, I have not hesitated to utilize symbolic notations, mathematical formulas, structural diagrams, or technical terminology where this has seemed necessary or appropriate. I have also sought to bear in mind Rawls’ wise observation that, in explicating commonsense morality, one must learn from one’s predecessors and recognize that “morals is not like physics: it is not a matter of ingenious discovery but of noticing lots of obvious things and keeping them all in reasonable balance at the same time” (Rawls 1951b: 579–580; cf. Kant 1993/1788: 8). Hence the ample use throughout the book of quotations, historical illustrations, parenthetical references, and other pedagogical resources. In this respect, my greatest ambition for the book will be realized if it can stimulate further research and serve as an accessible and useful resource for students of law, philosophy, and cognitive science to advance these fields of inquiry.
I have incurred enormous debts in writing this book and the articles and chapters from which it draws. Indeed, the list of individuals from whom I have received assistance in this regard is embarrassingly long. Some of these debts stretch back decades, and many of those to whom I am most indebted are sadly no longer alive.

I would like first to thank the four members of my dissertation committee under whose formal direction I worked on this project from 1995 to 2000: Noam Chomsky, Richard Miller, Jason Stanley, and Allen Wood. I learned a great deal from each of them, and each gave generously of their time in helping me to finish the dissertation. Noam and Allen, in particular, deserve special thanks for the extraordinary support and guidance they have given me over the course of nearly two decades.

During the early stages of my research, David Lyons, Terry Irwin, and Nicholas Sturgeon offered penetrating criticisms of my original dissertation proposals. I remain grateful for their assistance and hope they find some return on that investment here. I would also like to thank the other faculty members of the Sage School of Philosophy at Cornell University from whom I received my initial graduate training in philosophy: Richard Boyd, Mark Crimmins, Gail Fine, Harold Hodes, Karen Jones, and Sydney Shoemaker. During the early stages of this project, Carl Ginet, Scott MacDonald, Henry Shue, and Zoltan Gendler Szabo made time for me and offered useful advice. Finally, I wish to thank a terrific group of fellow graduate students at Cornell, with whom I had many stimulating conversations that sharpened my understanding of philosophy, and moral philosophy in particular. In this regard, I am conscious of specific debts to Tom Bennigson, Travis Butler, Rebecca Copenhaver, Jennifer Dworkin, Stephen Gardiner, Eric Hiddleston, Keith McPartland, Thaddeus Metz, Joe Moore, David Robb, Susanna Siegel, Chris Sturr, Christie Thomas, Martino Traxler, Ralph Wedgwood, and Jessica Wilson.

My first papers on Rawls were written for Christine Korsgaard and T.M. Scanlon, both of whom supplied me with helpful feedback. Chris and Tim were also instrumental in enabling me to participate in the Harvard Workshop in Moral and Political Philosophy from 1994 to 1996, where I presented my ideas on Rawls’ linguistic analogy for the first time. At Harvard, I also had the opportunity to take courses with Chris, Tim, Warren Goldfarb, Hilary Putnam, and Gisela Striker, and to engage in fruitful conversations about moral and legal theory with Anthony Appiah, Scott Brewer, Howard Gardner, Carol Gilligan, Duncan Kennedy, Fred Neuhouser, Robert Nozick, Derek Parfit, and Frederick Schauer. In June 1995, I had the pleasure of meeting John Rawls, and was lucky enough to discuss specific aspects of his work with him on several occasions thereafter. From 1995 to 1999, I served as a teaching Fellow for Seyla Benhabib, Cary Coglianese, and Kenneth Winston, and greatly benefited from conversations about ethics with all three. Finally, I received valuable input on this project from another exceptional group of fellow graduate students and visitors, including Carla Bagnoli, Sean Greenberg, Steven Gross, Erin Kelly, Thomas
Preface

Kelly, Rahul Kumar, Daniel Markovitz, Ian Proops, Faviola Rivera-Castro, Tamar Schapiro, Angela Smith, David Sussman, Alec Walen, and Leif Wenar. I thank all of these people for helping me to write this book and for making my tenure at Harvard such a rewarding experience.

In 1995, I took my first course in linguistics and cognitive science with Noam Chomsky. Shortly thereafter I became affiliated with MIT’s Department of Brain and Cognitive Sciences, and over the next four years I was fortunate to receive critical feedback and assistance on this project from many linguists and cognitive scientists, including James Blair, Sylvain Bromberger, Stephen Chorover, Danny Fox, Ted Gibson, Dan Grodner, Kenneth Hale, Morris Halle, Alan Hein, Nancy Kanwisher, Jerrold Katz, Frank Keil, Howard Lasnik, Alec Marantz, Gary Marcus, Jon Nissenbaum, Camillo Padoa-Schioppa, Steven Pinker, Mary Potter, Liina Pylkkänen, Whitman Richards, Javid Sadr, Roger Shepard, Mriganka Sur, Tessa Warren, Kenneth Wexler, and Yaoda Xu. While at MIT, I frequently attended graduate courses in linguistics taught by Michel DeGraff, Irene Heim, Kai von Fintel, Alex Marantz, and David Pesetsky. I thank them for giving me this valuable opportunity. Despite a busy schedule, Sally McConnell-Ginet generously agreed to sit as proxy for Noam Chomsky during my final dissertation examination. She also provided many helpful comments on the penultimate draft of that manuscript, for which I am very grateful.

I owe a special debt to Elizabeth Spelke, with whom I had the good fortune to collaborate during the early stages of this enterprise. Liz has taught me a tremendous amount about cognitive science, and I am extremely grateful for her guidance and enthusiasm. I would also like to emphasize my extensive debt to Cristina Sorrentino, with whom I was privileged to work closely for several years in launching the empirical component of this project. Other alums of the Spelke Lab to whom I owe thanks include Kirsten Condry, Sue Hespos, Lori Markson, Lori Santos, Bill Turkel, and Fei Xu. All of these individuals, especially Liz and Cristina, were influential in helping me to devise and analyze the trolley problems discussed in Part Two.

My debts to Philippe Schlenker, Chris Moore, Rajesh Kasturirangan, and, above all, Matthias Mahlmann and Joshua Tenenbaum, are even greater, and indeed would be hard to overestimate. For well over a decade, I have maintained a running dialogue with each of them about the ideas contained in this book. These delightfully stimulating conversations have had a profound impact on how I understand the fields of linguistics, cognitive science, neuroscience, and ethics. I extend to all five of them my heartfelt thanks for their advice, comments, criticisms, and, most importantly, their friendship.

Making the transition from graduate school to law school was not always easy, but Stanford Law School proved to be the ideal intellectual environment in which to do so. I was fortunate to find generous mentors at Stanford who cultivated my legal skills while also encouraging the research I had begun in graduate school. It would be difficult to overstate how much I owe Tom Grey,
xx  Preface

Pam Karlan, and Mark Kelman in this respect. For their support and encouragement, I would also like to thank Barbara Babcock, Joe Bankman, Richard Craswell, George Fisher, Barbara Fried, Hank Greely, Joe Grundfest, Larry Lessig, Mitch Polinsky, Jeff Strnad, Kathleen Sullivan, and Bob Weisberg. Among the friends and classmates whose patience I tested with my ramblings about the implications of cognitive science for jurisprudence, I am especially grateful to Fred Bloom, William Boyd, Kevin Collins, Cara Drinan, Kevin Driscoll, David Freeman Engstrom, Nora Freeman Engstrom, Roberto Gonzalez, Scott Hemphill, Josh Klein, Adam Kolber, Katherine McCarron, Sanjay Mody, Julian Davis Mortenson, John Ostergren, and Christian Turner.

After law school, I was fortunate to clerk for Judge Rosemary Barkett on the U.S. Court of Appeals for the Eleventh Circuit and to work alongside a great group of co-clerks: Cecily Baskir, Jeff Bowen, Keith Donoghue, Morgan Doran, Ian Elisasoph, Hannah Garry, Angela Littwin, and Alan Pegg. I thank Judge Barkett and my co-clerks for taking a keen interest in my academic work and offering their ideas and suggestions.

Georgetown University Law Center has been a remarkably warm and hospitable environment in which to teach, write, and pursue interdisciplinary scholarship. For their generous support of the research that went into completing the final stages of this book, I wish to thank Deans Alex Aleinikoff and Judy Areen and Associate Deans Larry Gottin, Vicki Jackson, and Robin West. David Luban, Henry Richardson, Nancy Sherman, and Robin West have been extraordinarily kind and generous philosophical mentors, and I could not ask for more enthusiastic and supportive colleagues than Randy Barnett, Julie Cohen, David Cole, Dan Ernst, Chai Feldblum, James Forman, Steve Goldberg, Lisa Heinzerling, Emma Coleman Jordan, Greg Klass, Don Langevoort, Marty Lederman, Carrie Menkel-Meadow, Naomi Mezely, Julie O'Sullivan, Nina Pillard, Mitt Regan, Mike Seidman, David Vladeck, Ethan Yale, and Kathy Zeiler. Many other colleagues at Georgetown have offered assistance or encouraged me in a variety of ways. In particular, I would like to thank Chuck Abernathy, Lama Abu-Odeh, Jane Aiken, Tom Beauchamp, Norman Birnbaum, Sue Bloch, Gregg Bloche, Rosa Brooks, Peter Byrne, Alisa Carse, Steve Cohen, Wayne Davis, Michael Diamond, Richard Diamond, Viet Dinh, Robert Drinan, Peter Edelman, Jim Feinerman, Heidi Li Feldman, Steven Goldblatt, Chuck Gustafson, Bob Haft, Paul Heck, Kris Henning, Craig Hoffman, Bryce Huebner, John Jackson, Neal Katyal, Pat King, Lauri Kohn, David Koplow, Mark Lance, Chuck Lawrence, Richard Lazarus, Amanda Leiter, Adam Levin, Judy Lichtenberg, David Lightfoot, Maggie Little, Wally Mlyniec, Jon Molot, Mark Murphy, Jim Oldham, Carol O’Neil, Les Orsy, Joe Page, Gary Peller, Terry Pinkard, Madison Powers, Nick Rosenkranz, Susan Deller Ross, Paul Rothstein, Peter Rubin, Alvaro Santos, Andy Schoenholtz, Jodi Short, Gerry Spann, Jane Stromseth, Peter Tague, Dan Tarullo, Jay Thomas, Philomila Tsoukala, Mark Tushnet, Rebecca Tusnhet, Carlos Vazquez, Pete Wales, and Si Wasserstrom.
The main arguments of this book have been presented at numerous workshops, seminars, conferences, and lectures, including legal theory workshops at Cardozo, Georgetown, Harvard, Illinois, Stanford, Virginia, and Yale law schools; philosophy talks at Cornell University, Dickinson College, Georgetown University, George Washington University, University of Berlin, University of British Columbia, University of New Mexico, and University of Toledo; graduate psychology seminars at Cornell University, MIT, and Stanford University; annual meetings of the American Association of Law Schools, American Philosophical Association, Cognitive Science Society, International Association for Philosophy of Law and Social Philosophy, Society for Philosophy and Psychology, and Society for Research in Child Development; philosophy and cognitive science conferences sponsored by the Air Force Office of Scientific Research, Brooklyn Law School, California Institute of Technology, Gruter Institute for Law and Behavioral Research, Harvard University Department of Psychology, Kellogg School of Management, MIT Computer Science and Artificial Intelligence Laboratory, Office of Naval Research, Princeton University Center for Human Values, Society for Evolutionary Analysis in Law, United Kingdom Arts and Humanities Research Council, University of Chicago Center for Law, Philosophy and Human Values, and Yale University Cowles Foundation for Economic Research; and invited lectures hosted by the MIT Culture and Technology Forum, MIT Department of Brain and Cognitive Sciences, Peter Wall Institute for Advanced Studies, Stanford University Department of Psychology, UCLA Center for Behavior, Evolution and Culture, and University of Maryland Cognitive Science Colloquium. I thank the organizers of these events as well as the many participants whose comments and criticisms helped to sharpen my understanding of various aspects of this research program.

This undertaking could not have been completed without additional support from a number of individuals and institutions. Joshua Greene supplied lengthy and detailed criticisms as a reviewer for Cambridge University Press and thereafter revealed his identity to initiate a fruitful exchange of ideas that continues to this day. I thank him along with a second, anonymous reviewer for their valuable input. Several other people read all or part of this manuscript in its early stages as a doctoral dissertation and gave me advice and encouragement, including Susan Dwyer, Gilbert Harman, Marc Hauser, Ray Jackendoff, James McGilvray, Julius Moravscik, Shaun Nichols, Jesse Prinz, Rebecca Saxe, Peter Singer, and M.B.E. Smith. For their outstanding editorial assistance, I wish to thank David Anderson, James Dunn, Catherine Fox, Helen Wheeler, and the other editors at Cambridge University Press, particularly Beatrice Rehl, who patiently shepherded this book to completion despite several unexpected delays. At various stages of this project, I relied on a talented group of research assistants, including Michael Dockery, Ben Dooling, Martin Hewett, Izzat Jarudi, Sean Kellem, Jennifer Rosenberg,
and Amber Smith. In addition, I have benefited enormously from the expert assistance of research librarians at Cornell, Georgetown, Harvard, MIT, and Stanford. I also would like to acknowledge the generous financial support of Georgetown University Law Center, Stanford Law School, Peter Wall Institute of Advanced Studies, and the Air Force Office of Scientific Research, along with the helpful feedback of the students in my Law and Philosophy, Law, Mind, and Brain, and Law, Cognitive Science, and Human Rights seminars. Finally, I wish to express my deepest appreciation to my siblings and extended family. They have been unfailingly supportive of this endeavor from start to finish.

My greatest intellectual debt in connection with this project is owed to Noam Chomsky. As a graduate student, visitor, and teaching Fellow in his courses on the foundations of linguistic theory, I heard him lecture many times throughout the 1990s. These lectures were among the most stimulating academic experiences of my life. They sparked a love of philosophy, linguistics, and cognitive science that has stayed with me ever since. Even more importantly, over the course of several decades I have spent countless hours in personal conversations with him, discussing not only academic subjects, but also history, politics, activism, and life in general. To say I am grateful for these opportunities and humbled by the interest he has shown in my work would be an extreme understatement.

Many other friends, colleagues, and associates from a variety of disciplines and professions have helped me to write this book by discussing its central ideas with me and offering their comments, criticisms, and suggestions. With apologies to those I may have omitted, I would like in particular to thank Tom Abowd, Samir Abu-Absi, Matt Adler, Ralph Adolphs, David Albert, Larry Alexander, Ronald Allen, Ernie Alleva, Scott Atran, Jack Balkin, Susan Bandes, Zenon Bankowski, Jonathan Baron, Charles Barzun, Clark Barrett, Dan Bartels, Daniel Batson, Paul Bello, Gregory Berns, Mira Bernstein, Paul Bloom, Jeremy Blumenthal, Susanna Blumenthal, Alexander Bolyanatz, Rob Boyd, Michael Brutman, David Brink, Russ Burnett, Colin Camerer, Susan Carey, Peter Carruthers, Bill Casebeer, Michael Chandler, Jules Coleman, Fiery Cushman, Antonio Damasio, Jonathan Dancy, John Darley, Stephen Darwall, Peter DeScioli, Frans De Waal, Peter Ditto, John Doris, Emmanuel Dupoux, Matti Eklund, Jakob Elster, David Faber, Dan Fessler, John Martin Fischer, Alan Fiske, Simon Fitzpatrick, John Flavell, Rob Folger, Ken Forbus, Charles Fried, Mark Geistfeld, Tamar Gendler, Tracey George, Itzhak Gilboa, Herbert Gintis, John Goldberg, Matthew Goldberg, Alvin Goldman, Oliver Goodenough, Noah Goodman, Ryan Goodman, Geoffrey Goodwin, Alison Gopnik, Tom Griffiths, Chris Guthrie, Knud Haakonsen, Jonathan Haidt, Todd Handy, Brian Hare, Paul Harris, Deborah Hellman, Tony Honore, Norbert Hornstein, Heidi Hurdl, Elisa Hurley, Douglas Husak, Pierre Jacob, Derek Jinks, Kent Johnson, Owen Jones, Craig Joseph, Paul Kahn, Frances Kamm, Jerry Kang, Robin Kar, Deborah Kelemen, Eldon Kerr, Melanie
Preface

Without doubt, I owe the most to my wife, Sarah; my children, Hannah and Andrew; and my parents, Ramzy and Maryse Mikhail. Words cannot express how much they mean to me. With love, devotion, and gratitude, I dedicate this book to them.

***