

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

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

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.

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

Humans use significant amounts of effort in assessing the appropriateness of their own and other individuals' behaviors. Moral emotions play a central role in both implicit and explicit moral appraisals, being an essential ingredient for human social cognition (Eisenberg, 2000; Fessler, 1999, 2001, 2004; Haidt 2003b). Nevertheless, the psychological structure and neural organization of moral emotions remain underspecified. Here we describe a framework that relates shared and distinctive phenomenological characteristics of moral emotions to a set of cognitive and emotional components. Based on clinical evidence and functional imaging studies (Beer, Heerey, Keltner, Scabini, & Knight, 2003; Eslinger & Damasio, 1985; Moll, de Oliveira-Souza, Bramati, & Grafman, 2002a; Moll, de Oliveira-Souza, Eslinger, Bramati, Mourao-Miranda, et al., 2002b), we suggest that moral emotions emerge as neural representations from the coactivation of brain regions that code for perception of social cues, event knowledge, and emotion (Moll, de Oliveira-Souza, & Eslinger, 2003). According to this hypothesis, the neural bases of moral emotion, knowledge, and attitudes are better explained by a representational approach, in contrast to the view of neural processes as guiding moral appraisals (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). This framework may provide a fertile ground for the development of a neuroscientific counterpart of a "theory of moral sentiments" as conceived by Adam Smith in the seventeenth century.

Background

The belief that human beings are endowed with a repertoire of emotions with a strong moral content is not new. In his classical oeuvre, Adam Smith (1723–1790) referred to them as moral sentiments (Smith, 1759/1976). Yet this idea has remained outside the reach of empirical science until quite

Box 1.1

Categories of moral emotions

Self-conscious emotions are linked to self-assessments (Tangney & Dearing, 2002). They arise from concerns about the opinions of others on self behaviors and on self-identity (Eisenberg, 2000). These emotions can be subdivided into **self-critical** (guilt, shame and embarrassment), which are linked to a sense of reduced social status or self-esteem, and **self-praising** (pride), which elicits a sense of increased self-esteem and social rank (Fessler, 2001).

Other-conscious emotions include those that are normally directed to others, and can be divided into three subcategories: **other-critical** (contempt/disgust and anger/indignation), **other-praising** (gratitude, awe), and **other-suffering** (pity or compassion). Other-critical moral emotions promote the punishment of others and the rupture of previous social links (Rozin, Lowery, Imada, & Haidt, 1999), while other-praising emotions drive positive reciprocity and cooperation. Other-suffering emotions play a key role in helping and altruistic behaviors (Haidt, 2003a).

recently. This is particularly true for normal and abnormal behaviors that take place within cultural contexts and which are heavily guided by subjective moral experiences and values (Moll, de Oliveira-Souza, & Eslinger, 2003). If, on the one hand neuroscience has grown up *pari passu* with the investigation of the cerebral substrates of cognitive architectures, on the other hand only in the past few decades have the neural underpinnings of social cognition and human emotions fallen within the reach of experimental science (Adolphs, 2003).

One important breakthrough in this respect has been the growing recognition that moral emotions comprise an essential ingredient of human social behavior. Moral emotions differ from basic emotions, such as fear and happiness, in that they are often linked to the interest or welfare either of society as a whole or at least of persons other than the agent (Haidt, 2003b). Although a final taxonomy of moral emotions has not as yet been completed (Eisenberg, 2000), they typically include (but are not restricted to) guilt, pity, embarrassment, shame, pride, awe, contempt, indignation, “moral” disgust, and gratitude (box 1.1 and table 1.1). As such, they are as instrumental in promoting the care of others, as well as cooperation and reciprocity, as in fostering blame, prejudice, and group dissolution (Moll et al., 2003; Schulkin, 2004). In this chapter, we will use the term “moral” according to its more ample definition—*moralis*, which refers to the laws of proper behavior and customs in daily life (Glare, 1982). In our view,

Table 1.1
Categories of moral emotions (subdivisions)

	Self-conscious		Other-conscious		
	<i>Self-critical</i>	<i>Self-praising</i>	<i>Other-critical</i>	<i>Other-praising</i>	<i>Other-suffering</i>
Guilt	•				
Shame	•				
Embarrassment	•				
Pride		•			
Indignation/anger			•		
Contempt/disgust			•		
Pity/compassion					•
Awe/elevation				•	
Gratitude				•	

Source: Modified from Haidt (2003b).

this broader definition provides a better approximation of what lay people consider appropriate or inappropriate in social conduct, in contrast to the more limited definitions employed by western moral philosophy in intellectual and professional circles.

In theory, moral emotions can be sorted into groups, or families, according to the particular type of stimuli or conditions that elicit them (elicitors) and by the specific action tendencies they evoke (Haidt, 2003b). However, these two basic classification criteria should not conceal their structural complexity. For example, although the moral emotions of shame and embarrassment share a general propensity to encourage “proper” behavior, especially in the presence of higher-ranking members of a social group, they can be distinguished in that embarrassment is felt when one violates social conventions, while shame is elicited by one’s own attribution of a reduced self-value following a violation of a moral norm (Tangney, 2000b). Such naturalistic observations led to three working hypotheses that have currently occupied a number of researchers.

In the first place, the aforementioned subtle distinctions between closely related moral emotions would require an equally sophisticated cognitive apparatus to decode the often fleeting and concealed social cues from the surrounding social milieu. These cognitive mechanisms are obviously more

complex than those employed in decoding basic emotions. Second, despite their neurobiological diversity and dependence on cultural codification, moral emotions have been reliably observed across individuals and societies (Fessler, 2004; Harris, 2003; Tangney & Dearing, 2002). However, what is experienced as a moral emotion is probably the result of the blending of elementary subjective emotional experiences, which are ubiquitous in mammals, with emotional and cognitive mechanisms that are typically human. The precise nature of these structural building blocks of moral emotions remains to be empirically determined and is the focus of this chapter. Third, the cognitive mechanisms and the subtle distinctions that prevail among the components of a given moral emotion, as well as among different families of moral emotions, must find an intimate correspondence to the neural representations that mediate each of them.

Moral Emotions and the Human Brain

Some of the most compelling data relating moral behavior, moral emotions, and the brain come from the remarkable overlap between the brain regions activated by explicit moral judgment tasks in functional magnetic resonance imaging (fMRI) studies and those related to disturbances of social behaviors in lesion studies (Eslinger & Damasio, 1985; Eslinger, Flaherty-Craig, & Benton, 2004; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Heekeren, Wartenburger, Schmidt, Schwintowski, & Villringer, 2003; Moll, de Oliveira-Souza, Bramati, & Grafman, 2002a; Moll, Eslinger, & Oliveira-Souza, 2001; Seger, Stone, & Keenan, 2004; Stone, Cosmides, Tooby, Kroll, & Knight, 2002; Takahashi, Yahata, Koeda, Matsuda, Asai, & Okubo, 2004). Most of these studies employed visual presentation of written material, either simple statements or short stories. Activated regions included the prefrontal cortex (PFC), encompassing the frontopolar (FPC) and orbitofrontal (OFC) cortices, the superior temporal sulcus (STS) region, and the anterior temporal cortex. Limbic regions, such as the amygdala and ventral striatum, have also been shown to be activated, though less consistently (probably owing to intrinsic technical limitations associated with functional magnetic resonance imaging).

The brain regions activated in moral judgment tasks have been implicated in experiencing emotion (amygdala), semantic memory (anterior temporal cortex), perception of social cues (STS region) and decision making—either quick and implicit, based on heuristics, or slow and reflective, based on prospective reasoning (OFC, FPC). Remarkably, a similar

brain network is engaged when normal subjects passively view visual stimuli with moral connotations (e.g., abandoned children, war scenes, interpersonal aggression), while brain regions associated with more basic emotional mechanisms (e.g., amygdala, brainstem) are activated both by unpleasant scenes without clear moral implications (e.g., dangerous animals, mutilated bodies) and by moral scenes (Moll et al., 2002b). We have postulated that the coactivation of PFC-temporolimbic networks corresponds to the cognitive-emotional states associated with “moral sensitivity.” Moral sensitivity provides a critical mechanism by which humans automatically attribute moral significance to ordinary events and behaviors (Moll et al., 2002b).

These findings show that the PFC and associated regions are spontaneously engaged whether or not decisions or behavioral outputs are required, thereby suggesting that the PFC is not merely manipulating information stored elsewhere, but is actively involved in representing social knowledge. This view stands in contrast to the dualistic notion that the PFC sustains executive processes, for example conflict monitoring or cognitive set shifting, and that these rational processes compete with emotional ones (Greene et al., 2004). Instead, according to our view, moral emotions would neither compete with rational processes during moral judgments, nor result from them. Most likely, moral emotions help guide moral judgments by attaching value to whichever behavioral options are contemplated during the tackling of a moral dilemma.

Evidence pointing to differences in the neural representation of individual moral emotions within the networks mentioned is starting to emerge. Pity, indignation, guilt, and embarrassment activate partially distinct cortical and subcortical brain regions in moral judgment tasks (Moll et al., 2003; Takahashi et al., 2004).

The Integration of Context-Dependent Attitudes with Moral Emotions

Psychologists, clinicians, and philosophers have long assumed that humans share cognitive processes, such as inductive and deductive reasoning, attention, conflict monitoring, and semantic categorization. Inferring cognitive and neural mechanisms from observed behaviors, however, can be misleading (Wilkinson & Halligan, 2004). This issue becomes even more problematic when cultural effects are at play because these are often complex and difficult to control experimentally. In other words, contextual embedding plays a central role in human cognition and emotion. For example, westerners and East Asians differ in their use of

logic versus dialectical principles and rules of categorization when they make causal attributions and predictions (Ji, Peng, & Nisbett, 2000; Nisbett & Masuda, 2003). Even if the engagement of a cognitive mechanism, say attention, is required to the same degree to solve a task, the content of the brain representations it helps elicit can be extremely variable.

Moral emotions are elicited in response to departures from implicit and explicit social norms and stereotypes, which code for individual attitudes and beliefs (Nichols, 2002b). The contextual elements linking moral emotions to norms can be quite variable and are to a great extent shaped by culture (Ehrlich, 2000). The PFC is a critical region for learning, storing, and binding social knowledge to contextual elements (Wood & Grafman, 2003). Damage to the anterior sectors of the PFC at an early age seems to prevent the acquisition of moral attitudes and values, leading to gross disturbances of social conduct (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Eslinger et al., 2004). In addition, ventral damage to the PFC impairs implicit stereotypes (Milne & Grafman, 2001). Social stereotypes and attitudes are associations of the self or other persons with negative or positive attributes (Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott, 2002). The word “implicit” means that attitudes may be hidden from public view and even from conscious awareness. Thus, even when they linger outside conscious awareness, attitudes are powerful determinants of behaviors that can be justified by convincing logico-verbal arguments.

Recent evidence (Cunningham, Nezlek, & Banaji, 2004) shows that attitudes toward a wide range of social issues, such as war or abortion, activate partially overlapping brain networks (including the amygdala, PFC, OFC, and anterior temporal cortex), regardless of whether such attitudes are evoked explicitly (subjects required to judge concepts as good or bad) or implicitly (concepts judged as concrete or abstract). It is interesting that activity in the anterior PFC was correlated with scorings of “ambivalence” toward attitudes. In accord with the role of the anterior PFC in prospective evaluations and thinking about the future (Goel, Grafman, Tajik, Gana, & Danto, 1997; Koechlin, Ody, & Kouneiher, 2003; Okuda, Fujii, Ohtake, Tsukiura, Tanji et al., 2003), we have suggested that this region also plays a role in a moral calculus (Moll et al., 2003; Moll, Eslinger, & Oliveira-Souza, 2001). A moral calculus results from the ability to envision a number of action-outcome options in a parallel fashion, and compare their relative weights. While Greene and colleagues have recently espoused this view (Greene et al., 2004), their interpretation differs from ours in a subtle but

important way: they hypothesize that the PFC performs a cognitive control function, inhibiting emotionally guided responses and leading to more rational moral choices (i.e., utilitarian outcomes). In contrast, we maintain that the anterior PFC (the FPC, in particular) represents different aspects of social knowledge, which are bound to emotional relevance. These representations would then guide the assessment of social-emotional outcomes associated with behavioral choices.

Moral Emotions from a Cognitive Neuroscientific Perspective

Our approach aims to identify some of the building blocks of moral emotions from a neurobiologically useful perspective. By “useful” we mean a coherent scenario that will ultimately guide the design of experiments to probe the neural mechanisms of human social behavior, both in normal individuals and in selected patients with strategic cerebral lesions. A sound framework should be able to predict the occurrence of each moral emotion from different combinations of basic components. We propose a scheme with six main components that seeks to provide such predictions. These components were defined on the basis of psychological and neurobiological plausibility. The first three (attachment, aggressiveness, and social rank/dominance/self-esteem) correspond to basic cognitive and emotional mechanisms that are widely represented across species. The fourth and fifth (outcome assessment and agency or intentionality) rely on more differentiated integrative systems and have been extensively studied in humans in recent years. The sixth (norm violation) has been attributed to the role of the PFC in sequential and motivationally relevant knowledge of an event (Wood & Grafman, 2003).

We describe each component here and briefly address how they might relate to the moral emotions. Then we address how each moral emotion may emerge from specific combinations of these components.

Attachment

Attachment provides the basic ingredient for interindividual bonding and affiliative behaviors, such as mother-offspring ties. Subcortical and limbic structures, including the ventral striatum, septal nuclei, amygdala, and hypothalamus, as well as hypothalamic-hypophyseal hormones such as oxytocin, have been implicated in attachment (Bartels & Zeki, 2004; Insel & Fernald, 2004; Keverne & Curley, 2004). While its essential aspects rely on primitive neural systems, attachment is generally bound to social cognitive mechanisms, including perception of social cues and inference of the

mental states of others. This combination provides the basis for emotional empathy, which is an affective response that stems from understanding another's emotional state, so that one has feelings similar to those of the other person. This ability arises during development, leading to the differentiation between one's own and others' internal states, and allows the emergence of pity and compassion (Eisenberg, 2000). We hypothesize that attachment is an essential ingredient for the emergence of guilt and gratitude. Guilt is often elicited in situations in which one feels bad for causing harm to another person, to other creatures (including animals and plants), and even to objects and abstract values to which one is emotionally attached. Gratitude in turn is associated with strengthened social bonds and an inclination to repay favors to benefactors, and thus might depend on basic mechanisms of attachment.

Aggressiveness

Experimental research has extensively shown that aggression occurs in disputes regarding sex, territory, and feeding. Dopamine has been implicated in the processing of signals of aggression in social-agonistic encounters in several species, including humans. Accordingly, D2-class receptor dopaminergic antagonism leads to a selective disruption in anger recognition (Lawrence, Calder, McGowan, & Grasby, 2002). Basic aggressiveness, combined with sophisticated social cognitive mechanisms of reputation assessment and representation and violations of social norms, is an essential ingredient of instrumental and moralistic aggression (Arsenio & Lemerise, 2004). The subjective experience of other-critical emotions, such as disgust, contempt, and indignation, is probably dependent on the appropriate functioning of neural circuits supporting aggression. These circuits include the amygdala, septal area, hypothalamus, and cingulate cortex, as well as their temporal and frontal connections (Mega, Cummings, Salloway, & Malloy, 1997; Moll, de Oliveira-Souza, Tovar-Moll, Ignacio, Bramati, et al., 2005a; Volavka, 1999). Although there is a close link between aggressiveness and dominance, they may be dissociable.

Social Rank/Dominance/Self-Esteem

Social animals are organized into hierarchical structures. Displays of submission (decreased eye contact and apparent body size) and dominance (increased eye contact and apparent body size) are widespread in mammals (Haidt, 2003b). Primates possess highly structured dominance hierarchies that regulate access to food resources, mating, and other social privileges (de Waal, 1996). Social status marks one as a good or poor partner for

future interactions. There is extensive evidence pointing to the role of dopaminergic and serotonergic pathways in social dominance (Morgan, Grant, Gage, Mach, Kaplan et al., 2002). Though both enhanced dopaminergic and serotonergic action have been related to increased dominance, these neurochemical systems probably exert partially separable effects. Increased serotonergic activity provided by selective serotonin reuptake inhibitors (SSRI) has been linked to a decrease in harm avoidance and hostility and an increase in dominance in social encounters (Brody, Saxena, Fairbanks, Alborzian, Demaree et al., 2000). Serotonin most likely exerts a modulatory effect on social interactions that depend on social status (Edwards & Kravitz, 1997).

Further support for the dissociation between aggressiveness and dominance and self-esteem comes from the observation that higher levels of serotonin promote constructive social interactions by decreasing aggression (Young & Leyton, 2002). Self-esteem is a more complex notion that has been defined as the association of a self-concept with emotional valence (Greenwald et al., 2002). Although these dopaminergic and serotonergic pathways are difficult to circumscribe anatomically, brainstem regions (e.g., the midbrain ventral tegmental area), basal forebrain nuclei, and their projections to the OFC and the subgenual area of the anterior cingulate cortex are probably critical for these mechanisms.

Fessler (2004) has suggested that proto-forms of pride and shame developed to motivate the quest for social dominance, but were refashioned into a new class of emotions that allow one to sense the subjective experiences of other individuals toward oneself, thus extending dominance-striving motives into the cultural world of humans. Shame and pride are chiefly linked to social dominance and self-esteem and help regulate approval seeking and dominance striving. They can facilitate cooperation by physiologically enabling internally motivated states that enhance conformity to social norms (Fessler, 2004). Accordingly, self-esteem is influenced by serotonin metabolism, which regulates the degree of engagement of shame and pride (Fessler, 2001). According to the present framework, pride, shame, and embarrassment require the binding of social cognitive mechanisms (e.g., agency and evaluation of the mental states of others) with core motivational states. Some of these sophisticated social cognitive abilities have been described as “theory of mind,” “perspective taking” and “simulation theory” (Gusnard, Akbudak, Shulman, & Raichle, 2001; Happé, 2003; Saxe, Carey, & Kanwisher, 2004a). According to the formulation presented in table 1.2, perspective taking supports the ability to perceive that one is being observed, which plays an important role in the elicitation of pride, shame, and embarrassment.

Outcome Assessment

The ability to predict long-term outcomes is uniquely developed in humans and is tightly linked to the development of the PFC circuits. Different functions have been ascribed to the PFC: attention control; monitoring; and adjustment of task scheduling, switching, and planning. A common denominator is the role of the PFC in dealing with events. The structured event complex (SEC) framework (Wood & Grafman, 2003) provides a unifying mechanism to explain how the PFC stores information in event complexes through sequential binding and integrates contextual information under the influence of factors such as familiarity and predictability. In certain situations, moral emotions are tightly linked to the ability to assess outcomes. When humans contemplate a behavioral option linked to short-term rewards, they often estimate (implicitly or explicitly) the possible desirable or undesirable long-term outcomes of that action. This ability is ascribed to the anterior PFC and its role in representing structured event complexes with a long duration (Wood & Grafman, 2003) and in assigning values to future rewards (Tanaka, Doya, Okada, Ueda, Okamoto, & Yamawaki, 2004). When such predictions take place in social contexts, they can influence the occurrence of certain moral emotions prior to or after carrying out an action (e.g., one considers attaining a selfish, short-term reward, but feels guilty and refrains from acting upon foreseeing that the act would make another person suffer; or one acts in the first place but then realizes the possibility of adverse consequences to others).

Agency and Intentionality

Humans often attribute agency and intentionality to others. This is a pervasive cognitive mechanism that allows one to predict the actions of others based on spatiotemporal configurations and mechanical inferences, and on their putative internal states, goals, and motives. Although a high degree of overlap may exist between attributing agency and intentionality, they are not the same thing. If you bump into an object and it falls and breaks, you are the agent of a mechanical action. But if you throw it against a wall, in addition to being the agent, you performed an intentional action. Agency and intentionality can hardly be considered as basic modules of the human mind, however. These abilities draw on evolved brain circuits that are able to link current appraisals of actions and motivational states of the self or of another person to ongoing events.

Brain regions that have been implicated in agency and intentionality include the parietal cortex, the insula, and the motor cortex, as well as the

Table 1.2
 Predicted relationships between moral emotions and putative neurocognitive components

	Attachment	Aggressiveness	Social dominance attribution				Outcome assessment				Intentional action (self)	Intentional action (other)	Agency		Social norm violation	
			Self by another	Self by oneself	Other by oneself	Being observed by another	Bad, self	Bad, other	Good, self	Good, other			Self	Other		
Guilt	•			→						•				•		•*
Shame			→	→						•				•		•
Embarrassment			→	→						•				•*		•
Pride			→	→										•*		•
Indignation/anger		•								•*						•
Contempt/disgust		•		→												•*
Pity/compassion	•															
Awe/elevation	•			→												•
Gratitude	•			→												•

Note: Black dots (•) indicate a key role of the cognitive variable in the genesis of the corresponding moral emotion. Arrows indicate increased (↑) or decreased (↓) attribution of dominance or social status, according to the self and other perspectives (e.g., decreased “self by another” means that one believes that his or her own social status is judged to be reduced from the point of view of another person). Asterisks indicate that the role of the component for a given moral emotion can be either variable or only modulatory (but not essential).

medial PFC and the STS region (Daprati, Nico, Franck, & Sirigu, 2003; Jiang, Saxe, & Kanwisher, 2004; Pridham, Saxe, & Limbo, 2004; Saxe, 2004; Saxe et al., 2004a; Saxe & Kanwisher, 2003; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004b; Sirigu, Daprati, Ciancia, Giraux, Nighoghossian et al., 2004; Sirigu, Daprati, Pradat-Diehl, Franck, & Jeannerod, 1999). Because both agency and intentionality can independently provide sharp discriminations among certain moral emotions, they were included as separate subcomponents in table 1.2, although they probably partially share neural representations. For example, feeling guilt generally requires recognition of oneself as the agent of an action that leads to a bad outcome to another person; intentionality here is not as critical as agency. If the same outcome is contemplated but one does not perceive oneself as the agent, pity may be experienced instead. Intentionality plays a more central role than agency for other moral emotions; indignation, for example, tends to be more pervasive when there is a voluntary violation of social norms.

Norm Violation

Norms are abstract concepts that are firmly encoded in the human mind. They differ from most kinds of abstractions because they code for behavioral standards and expectations and are often associated with emotional reactions when violated (Nichols, 2002b). Despite the large variation in the content of norms and in how contextual elements are bound to them, virtually all human societies rely on normative systems for setting common grounds and limits for interpersonal transactions. A violation of a norm signals the violation of expectations and preferences of other people in the social milieu. Norm violations are salient and conspicuous events that draw upon the cultural specifics of a social group and their context. A violation of social norms is a critical element in the elicitation of certain moral emotions, although it is not sufficient for specifying which moral emotion will be elicited. For example, in the oddball task, there is a predictable chain of events (e.g., arbitrary sensory stimuli of a given category) that is suddenly broken by a nonmatch, unexpected event, or stimulus. This oddball effect, which marks a violation of an arbitrary norm, has been demonstrated to elicit brain responses in regions linked to conflict monitoring, behavioral flexibility, and social response reversals, such as the anterior cingulate cortex, anterior insula, and the lateral OFC (Blair & Cipolotti, 2000; Hornak, Bramham, Rolls, Morris, O'Doherty, et al., 2003). If arbitrary events were replaced by ones associated with social attitudes, however, violation of expectancies would be encoded as violations of social norms. Indeed, these brain regions were shown to be activated by written

stimuli describing moral violations that evoke anger and disgust (Moll et al., 2005a).

Knowing if and who violated a social norm requires both detecting the violation and attributing agency. These ingredients are necessary for eliciting and discriminating between self-critical (guilt, shame, and embarrassment) and other-critical (contempt, disgust, and indignation or anger) moral emotions. Guilt and shame are associated with self-violations of a social norm, while contempt, disgust, and anger or indignation result from the recognition that another person carried out the transgression. One still controversial issue is whether shame and embarrassment can be differentiated on the basis of the severity of the social violation, or on a categorical distinction between moral violations and violations of social conventions (Eisenberg, 2000; Harris, 2003; Tangney, 2000b).

Predicting Moral Emotions on the Basis of Cognitive-Emotional Components

According to our framework, the elicitation of moral emotions is implemented by dynamic PFC-temporolimbic network representations, which arise from coactivation of the components described earlier. Table 1.2 illustrates how these components might combine to give rise to specific moral emotions. Here we briefly describe the “recipe” of each moral emotion according to the postulated neurobiological ingredients and cite the brain regions putatively involved in each. It should be emphasized, however, that at least in some cases the evidence is still preliminary and should be considered provisional.

Guilt

Guilt emerges prototypically from (1) recognizing or envisioning a bad outcome to another person, (2) attributing the agency of such an outcome to oneself, and (3) being attached to the damaged person (or abstract value). Depending on the circumstance, the recognition that (4) a social norm was violated may lead to increased guilt. Guilt is associated with reduced self-esteem, which is not necessarily dependent on being observed by another person (Eisenberg, 2000; Harris, 2003; Tangney, 2000a, 2002; Tangney & Dearing, 2002). Recent neuroimaging data showed the involvement of the anterior PFC, the anterior temporal cortex, the insula, the anterior cingulate cortex, and the STS region in guilt experience (Shin, Dougherty, Orr, Pitman, Lasko et al., 2000; Takahashi et al., 2004).

Shame

This emotion follows (1) an action attributed to self-agency that is associated with (2) a violation of a social norm, leading to (3) a bad outcome to oneself (at least indirectly, such as damage to one's reputation), and (4) reduced self-esteem (and social dominance) as judged by oneself and others, which is often reliant on (5) the other's awareness of one's actions (Eisenberg, 2000; Harris, 2003; Tangney, 2000b), although this aspect is questioned by other authors (Baumeister, Stillwell, & Heatherton, 1994). While there are no available brain imaging studies on shame, brain regions similar to those demonstrated for embarrassment should be involved. The ventral part of the anterior cingulate cortex, also known as the subgenual area, has been associated with depressive symptoms, which are linked to decreased self-esteem (Fu, Williams, Cleare, Brammer, Walsh et al., 2004). Therefore, this brain region might also play a more specific role in the neural representation of shame.

Embarrassment

Embarrassment has traditionally been viewed as a variant of shame (Lewis & Steiben, 2004). In egalitarian cultures that have an independent construal of the self, embarrassment splits off from shame (Eisenberg, 2000; Haidt, 2003b; Tangney, 2002). Embarrassment follows (1) being the agent, (2) violating less severe social norms or conventions compared with shame, leading to recognition of (3) a bad outcome to oneself, and (4) a non-pervasive reduction of one's social dominance, which is dependent on (5) others' awareness (being observed). In contrast to guilt and shame, however, embarrassment is not thought to be associated with a significant reduction of self-esteem or a sustained decrease in self-attributed social dominance. The neural correlates of embarrassment have recently been addressed (Berthoz, Armony, Blair, & Dolan, 2002; Takahashi et al., 2004) and include the medial PFC, the anterior temporal cortex, the STS region, and the lateral division of the OFC.

Pride

The polar opposite of shame and embarrassment is pride. It is associated with (1) being the agent of an action leading to (2) good outcomes for oneself, in which there is an (3) increased attribution of social dominance and self-esteem. Perhaps not mandatory for the experience of pride is (4) the observation of one's socially desirable actions by others (being observed) and (5) performing an intentional action. So far, there is no clear evidence for the neural representation of pride, although it has been shown that

patients with OFC lesions may experience this emotion inappropriately (Beer et al., 2003). We hypothesize that brain regions involved with mapping the intentions of other persons (e.g., the medial PFC and the STS region), and regions involved in reward responses (OFC, hypothalamus, septal nuclei, ventral striatum) may play a role in this emotion.

Indignation and Anger

Indignation and anger are elicited by (1) observing a norm violation in which (2) another person is the agent, especially if (3) the agent acted intentionally. Indignation relies on (4) engagement of aggressiveness, following an observation of (5) bad outcomes to the self or to a third party. We and others have shown that indignation evokes activation of the OFC (especially its lateral division), the anterior PFC, anterior insula, and anterior cingulate cortex (Blair, Morris, Frith, Perrett, & Dolan, 1999; Moll et al., 2005a).

Contempt

Contempt has been considered a blend of anger and disgust, and has sometimes been considered as a more subtle form of interpersonal disgust (Haidt, 2003b; Plutchik, 1980). For this reason it will be considered here together with disgust. Contempt and disgust mark distinctions of rank and prestige, especially in hierarchical societies. They are evoked by (1) devaluation of another's social status by oneself, i.e., attribution of decreased social dominance to another person (although upward contempt also exists), and (2) recruitment of aggressiveness to some degree. Contempt and disgust can be enhanced by (3) minor social norm violations in which (4) another person is the agent. Severe social violations, especially when intentional, will generally trigger indignation instead. While contempt has not been nominally addressed in experimental neuroanatomical studies, the neural representations of disgust have been shown to include the anterior insula, the anterior cingulate and temporal cortices, the basal ganglia, the amygdala, and the OFC (Buchanan, Tranel, & Adolphs, 2004; Calder, Keane, Manes, Antoun, & Young, 2000; Fitzgerald, Posse, Moore, Tancer, Nathan, & Phan, 2004; Moll et al., 2005a).

Pity and Compassion

These feelings are described as "being moved by another's suffering" (Haidt, 2003b; Harris, 2003). They require (1) the recognition of bad outcomes to another person (or social group) and (2) a sense of attachment. Preliminary functional imaging results in normal subjects point to the involvement of

the anterior PFC, the dorsolateral PFC, the OFC, the anterior insula, and the anterior temporal cortex in pity or compassion (Moll et al., 2003). Further studies with more refined imaging techniques may be able to test our prediction that certain limbic regions, such as the hypothalamus, septal nuclei, and ventral striatum, would also be involved.

Awe and Elevation

Awe and elevation are poorly understood emotions that have received more attention recently (Haidt, 2003a). While these emotions can be experienced in nonmoral situations (e.g., when admiring a beautiful landscape or artwork), in social contexts they are likely to be elicited when one (1) observes the action of another agent and (2) experiences a positive feeling of attachment to this agent or to its accomplishment. Often this will be linked to (3) good outcomes to a third party (i.e., an altruistic action), although in other instances it can derive from observing other kinds of highly praised acts (e.g., an act of courage or an act that springs from other outstanding skills). Finally, awe is associated with the attribution of (4) increased social status to another agent. While the neuroanatomy of awe is still obscure, we predict that it will involve limbic regions associated with reward mechanisms, including the hypothalamus, ventral striatum, and medial OFC (Rolls, Kringelbach, & de Araujo, 2003; Tanaka et al., 2004), as well as cortical regions linked to perspective taking and perceiving social cues, such as the anterior PFC and the STS region (Moll et al., 2003).

Gratitude

This emotion is elicited by (1) detecting a good outcome to oneself, attributed to (2) the agency of another person, (3) who acted in an intentional manner to achieve the outcome. (4) Gratitude is associated with a feeling of attachment to the other agent and often promotes the reciprocation of favors. Recent studies using game-theoretic methods have started to address the brain regions involved with positive reciprocity (retribution of favors), with obvious implications for the neural underpinnings of gratitude. Activated brain regions included the ventral striatum, the OFC, and the anterior cingulate cortex (Rilling, Gutman, Zeh, Pagnoni, Berns, & Kilts, 2002; Singer, Kiebel, Winston, Dolan, & Frith, 2004).

Final Remarks and Future Directions

We have presented a novel approach to explain the shared and distinctive phenomenological characteristics of moral emotions on the basis of a set

of social cognitive and emotional component representations. The basic cognitive components postulated can be linked to distinct neural systems based on evidence from neuroanatomical, neurochemical, lesion, or functional imaging studies. Owing to the complexity of this endeavor, this initial approach cannot be interpreted as a complete or inclusive model. Especially with respect to the representation of human social dominance and attitudes, and their relationships to self-esteem, the neuroscientific evidence is preliminary.

Future functional imaging studies and clinical investigations in patients with focal brain lesions aiming at the neural representation of moral emotions should be able to independently manipulate the components outlined here to test predictions inspired by this framework. Also, it would be of great interest to explore if and how these components interact during development in the path to moral maturity. It was not within the scope of this work to comprehensively describe the more general implications of moral emotions for human social cognition. However, we believe that a more detailed neuroscientific theory of moral emotions and attitudes is urgently needed to explain a variety of uniquely human aspects of social cognition, such as the culturally and individually shaped context dependence of human social behavior.

Moral emotions might prove to be a key venue for understanding how phylogenetically old neural systems, such as the limbic system, were integrated with brain regions more recently shaped by evolution, such as the anterior PFC, to produce moral judgment, reasoning, and behavior. The pace of development of the new field of moral neuroscience will critically depend on the free exchange of ideas, open and unbiased scientific discussions, and the design of experiments and models that link the humanities and biological sciences.