14

The Disunity of Morality

Walter Sinnott-Armstrong

ONE OF THE most important lessons from the first decade of research in moral neuroscience is that morality is not unified in the brain or anywhere else. This lesson should shape the next decade of research into moral neuroscience. However, many neuroscientists, psychologists, philosophers, and others still assume that all moral judgments share some common and distinctive neural basis.

Some neuroscientists say so explicitly. For example, in a review Moll et al. (2005, 799) write, "Recent functional imaging and clinical evidence indicates that a remarkably consistent network of brain regions is involved in moral cognition." Notice the lack of qualification. Moll and colleagues do not say that this network is involved in *some* moral cognition. Instead, they suggest that the same network of brain regions is involved in *all* moral cognition. Conversely, although Moll and colleagues can admit that each individual brain region in the identified network is also involved in some nonmoral cognition, it would not be interesting to find a network of brain regions involved in moral cognition if that whole network is also involved in many kinds of *nonmoral* cognition. If the network subserves all judgments—whether moral or not—then it might show something about judgments without showing anything special about morality. Thus, Moll and colleagues seem to suggest that the discovered network of brain regions is distinctively moral.

Similar claims are implied by psychologists who claim to discover "the essence of morality." Recently, Gray, Young, and Waitz wrote, "We investigate whether all moral judgments can be explained by appealing to a dyadic template" (2012, 103), and their answer is affirmative. Assuming physicalism, the proposed psychological template must have some neural basis. The psychological template would hardly be a single "essence" if its neural basis were not unified. So these psychologists seem to be committed to the claim that all moral judgments share a distinctive and unified neural basis.

Other moral scientists and philosophers are less explicit but still seem to assume that moral judgments are unified in the brain. For example, Greene and Haidt cite a small sample of moral judgments but draw conclusions about moral judgments in general: "Neuroimaging studies of moral judgment in normal adults, as well as studies of individuals exhibiting aberrant moral behavior, all point to the conclusion, embraced by the social intuitionist model, that emotion is a significant driving force in moral judgment" (2002, 522). Again, notice the lack of qualification, so they seem to be referring to all moral judgments. A generalization from *some* moral judgments to *all* moral judgments can be justified only if moral judgments are unified by some feature that supports those generalizations. Otherwise, how could they rule out the hypothesis that emotion is a significant driving force in some but not all moral judgments? Thus, regardless of whether they are aware of it, scientists who generalize from small samples to universal conclusions must be assuming that moral judgments are unified.

Philosophers notoriously share this assumption. Philosophical moral theories from utilitarianism to Kantianism to virtue theories all claim to apply to morality as such. Even Ross (1930), who postulates seven different prima facie moral duties, assumes that these seven duties share something distinctive that makes them moral duties instead of nonmoral duties. The same assumption is made by metaethical theories of the meanings of moral terms, the epistemic justifiability of moral judgments, and the relation between moral belief and motivation. These philosophical theories almost never admit the possibility that different moral theories might work for different kinds of moral judgments. Thus, although most philosophers do not talk about the brain directly, they seem to assume that moral judgments share some distinctive essence, and again that essence would seem to have some unified neural basis, assuming physicalism.

Admittedly, some philosophers, psychologists, and neuroscientists do not share this assumption. Flanagan (1991, 15–20) and Stich (2006, 181–189) explicitly reject the unity of morality. So do Greene and Haidt in their final conclusion: "There is no specifically moral part of the brain" (2002, 522) and "Morality is probably not a 'natural kind' in the brain" (2002, 523). I agree, and I am indebted to these predecessors.

Unfortunately, however, this lesson is not appreciated widely enough. Many philosophers, psychologists, and neuroscientists still do not agree—or at least they seem not to agree, because they continue to write, theorize, and design experiments as if morality is unified and, more specifically, as if there is some distinctively moral part of the brain. Moreover, even when careful philosophers and scientists add appropriate qualifications, there still remains a significant temptation for nonexpert commentators to overinterpret careful conclusions, such as when the popular media claim that scientists have found "the moral compass in the brain" (Sundby 2010). Claims like these stand in the way of progress.

My goal here is to convince these recalcitrant overgeneralizers to stop. I want to explain why and how morality fails to be unified. Of course, I cannot refute every possible proposal for unifying morality, so I won't try to do that. Instead, I will focus

on three main levels at which morality is often taken to be unified. First, I will discuss the content of morality, because variations in content display the challenge for my opponents. Second, I will discuss whether morality is unified at a neural level, because this volume focuses on the brain. Third, I will explore the function of morality, because the most common and popular proposals for unifying morality cite some function. Admittedly, even though these three proposals fail, some other candidate still might succeed in unifying morality.[1] Nonetheless, in the absence of another plausible proposal, I will conclude that nothing unifies morality. But, before I can assess any candidate for unifying morality, I need to clarify the standards that such candidates must meet.

14.1. What Is the Issue?

The question is basically whether morality is like memory. Once upon a time, philosophers and psychologists believed that memory is monolithic. Now memory is understood as a group of distinct phenomena that need to be studied separately (Tulving 2000). Memory includes not only semantic or declarative memory, such as remembering that a bat is a mammal, but also episodic memory, such as remembering seeing a bat yesterday. Memories can also be long-term or short-term (or working) memory, and procedural memory includes remembering how to do things, such as how to ride a bike. Thus, there are many kinds of memory, and they are not unified by any common and distinctive feature. They are not even all about the past, since you can also remember timeless truths, such as that pi is 3.14159 . . ., and you can also remember that you have a meeting tomorrow, even if you do not remember setting up the meeting or even who set it up. These kinds of memory differ not only in their psychological profiles and functions but also in their neural basis, as shown by both fMRI and by patients, such as H. M., whose brain lesions left him with severely impaired episodic memory but largely intact procedural and semantic memory. Such findings led most experts to accept that memory is not unified.

This recognition enabled progress. Neuroscientists could never find a neural basis for memory as such while they lumped together all kinds of memory. Psychologists could never formulate reliable generalizations about memory as long as they failed to distinguish kinds of memories. And philosophers could never settle how memory is justified if they conflated remembering facts and remembering how to ride a bicycle. Although these problems remain hard, progress became easier after recognizing that memory is not a single natural kind.

My thesis is that morality is like memory. Neither of them is unified, and admitting disunity makes progress possible in both areas. Moral neuroscience, psychology, and philosophy will become much more precise and productive if they give up the assumption that moral judgments all share a distinctive essence.

To make this thesis more precise, I first need to specify which class of things my thesis is about—the class that is claimed to be unified or not. The issue here is not about which moral judgments are true or justified or which judgments I (or you) endorse. Instead, the issue is only about which judgments are intended or supposed to be about moral norms as opposed to other kinds of norms, including aesthetic, prudential, economic, legal, conventional, or religious norms. People do distinguish kinds of norms and invoke one kind as opposed to another, but that practice by itself does not show that there are any real differences among those kinds of norms. Just as people might distinguish races even if races are not really distinct biologically, so people might distinguish kinds of norms even if there is no real difference between moral and nonmoral norms. The question here is about real rather than merely apparent differences. Nonetheless, we need to start with the norms that people classify as moral in order to determine whether or not anything really unifies those norms and distinguishes them from so-called nonmoral norms. Hence, I will count a judgment as a *moral* judgment if and only if people who make that judgment do or would think of it as similar enough in important respects to judgments that they take to be exemplars or paradigms of moral judgments. To sidestep controversy, I will focus on judgments that millions of people accept as moral judgments in this way, even if many people reject those moral judgments as false.

Critics might object that I am talking only about the judgments that millions of people (perhaps falsely) count as moral judgments rather than the judgments that really are moral judgments. To see why this objection is misdirected, compare races again. A biologist who wants to determine whether a race is unified biologically needs to start by specifying which people are supposed to be in that race and then asking whether anything that is biologically important is shared by and distinctive of those people. If a critic replies that the biologist who denies the unity of race is looking at people who are not really members of that race, then the critic assumes that there really is a race for people to be a member of. Analogously, critics who respond to my definition by saying that genuinely moral judgments are unified are assuming that there really is a class of genuinely moral norms for judgments to be a member of. That might be right, but we cannot assume it at the start. In order to avoid assuming either that moral judgments are unified or that they are not unified, we need some independent test of which judgments are moral judgments. That is why I appeal to the judgments that millions classify as moral.

Next, I also need to specify what it is for this class to be unified. I will call a group of phenomena *unified* if and only if some single feature that enables useful theoretical generalizations is present in all and only members of that group. Although this requirement is common for definitions, some critics might see it as too stringent: all and only! An alternative account of unity could be statistical, so that a group is unified when its members are more likely to share a trait with other members than with nonmembers. However, this statistical notion is too weak to serve important purposes. If moral judgments are only more likely to have certain features, then the fact that a new judgment has those features cannot show that the new judgment is a moral judgment, and the fact

that it lacks that feature cannot show that the new judgment is not a moral judgment. Moreover, if the relation is purely statistical, the feature need not capture what makes a judgment moral. Even if most (or nearly all) cities have a certain sort of public transportation system, and even if most areas with that sort of public transportation system are cities, the public transportation system is still not what makes this area a city. Similarly, a proposed account of what makes moral judgments moral must meet the more stringent test of fitting all and only moral judgments—or at least all and only judgments that millions of people see as moral judgments—in order to provide a test of which judgments are moral and in order to uncover what it is that makes them moral judgments.

My thesis, then, is that judgments that are intended to be about morality rather than some other kind of norm are not unified by any single common and distinctive feature that enables important generalizations about distinctive properties of those judgments. This thesis is compatible with the claim that moral judgments fall within a broader category that is unified. Even if moral judgments are not unified, for example, normative judgments still might be unified, perhaps because they (and only they) have some conceptual relation to reasons or motivation. My thesis is also compatible with unity within some narrower kinds of moral judgments, such as moral judgments about harm or the moral judgments of a certain person, society, or group. My thesis here is only that a particular level of generality—moral judgment—is not unified by any single important feature.

14.2. Is Morality Unified by Its Content?

The most obvious candidate for a unifying feature is content: what moral judgments are about. The idea cannot be simply that moral judgments are about morality, since that would beg the question of whether morality is a unified topic to be about.

Moral judgments are also not unified merely by being about what is right and wrong. DeScioli and Kurzban (2009, 282) say, "This paper examines 'morality,' meaning phenomena surrounding the concepts 'right' and 'wrong.'" Whatever they meant by this, the terms "right" and "wrong" are much too general to give the meaning of morality in particular, because many kinds of rightness and wrongness are independent of morality. There are right and wrong answers on history and math tests as well as right and wrong ways to bake a brioche or build a bridge. Thus, it is wrong (though not immoral) to identify morality with what is right and wrong. Moral judgments are moral because they judge acts to be right or wrong or good or bad in special ways, and those special ways are what need to be defined here.

A more plausible proposal about content claims that all and only moral judgments are about harm to others (Gert 2005; Mill [1859] 2011). In this vein, Grey, Young, and Waitz (2012, 16) write, "On our account, perceived suffering is not a distinct moral domain, but a core feature of all immoral acts." However, some moral judgments are about goods instead of harms. Imagine that a colleague gets tenure; and she will not suffer or be

harmed if you fail to give her a celebratory gift, partly because she does not expect any gift, and she is happy already. Nonetheless, many people would still judge that giving her a gift is nice or ideal, and the way in which it is nice is moral as opposed to legal, aesthetic, religious, conventional, or prudential.

Some theorists still might contrast moral ideals like niceness with moral requirements and prohibitions and claim that all moral requirements and prohibitions are about harm to others. Notice that this move already admits that harm to others cannot unify *all* moral judgments, since moral judgments do include judgments about moral ideals. Moreover, harm to others also cannot succeed in unifying even only the moral judgments that are about requirements or prohibitions. This is evident from a quick survey of moral requirements and prohibitions (which expands on Haidt 2011, which builds on Shweder et al. 1997):

Harm:
- *Death*: Do not kill.
- *Disability*: Do not blind, paralyze, or maim.
- *Loss of property*: Do not steal.
- *Physical pain*: Do not torture.
- *Psychological pain*: Do not insult or make people feel bad.

Justice:
- *Retributive*: Do not punish more or less than is deserved.
- *Distributive*: Do not treat people unequally.
- *Procedural*: Give everyone a fair hearing and a fair chance.

Dishonesty:
- Do not lie (or deceive).
- Do not break promises.
- Do not cheat (e.g., in games or in marriage).

Social position:
- *Hierarchy*: Do not disrespect or disobey your parents or elders.
- *Role*: Do your job and duty (e.g., as employee, citizen, or club member).
- *Loyalty* (to an in-group): Be patriotic. Don't rat on friends.

Purity:
- *Sexual*: Do not commit incest or necrophilia.
- *Gustatory*: Do not commit cannibalism.

These categories are vague, incomplete, and overlapping, and these rules need to be qualified. Still, the point here is only to illustrate some of the variety.

The first group of moral prohibitions explicitly concerns harm to others. So does retributive justice, insofar as punishment involves harm. In contrast, harm is not essential

to distributive and procedural justice, because unequal tenure gifts as well as procedures for awarding grants or prizes can be unjust apart from harm. Dishonest acts also seem morally forbidden even when they do not cause harms, as in the classic case of deathbed promises (unless we count false beliefs themselves as harms).

The disconnect between harm and moral judgments is even clearer in prohibitions based on social position. Sometimes nobody is harmed when one disobeys one's parents or the law, and yet many common folk and philosophers believe that we always have some (pro tanto) moral obligation to obey our parents and the law. Similarly, burning a flag is often seen as immoral because unpatriotic even when done secretly so that nobody is affected at all. Even when harmless, such acts are still judged as morally wrong by many people who accept such rules of hierarchy, role, and loyalty.

The same point applies to purity. If one eats part of a dead human body in secret when no disease is transferred, no habit begins, and the person died by accident, then cannibalism causes no harm. People might stretch to find some indirect harm, believed harm, or risk of harm, but that tendency does not show that such a tenuous relation to harm is what really makes people judge such acts to be immoral or what makes them classify such judgments as "moral" (cf. Haidt et al. 1993).

Many liberals and libertarians see the moral judgments that are not based on harm as irrelevant to real morality (Haidt and Graham 2007). However, what they mean by this claim seems to be that the people who make those judgments are mistaken. When people say that harmless impurity (allegedly in homosexuality, incest, and cannibalism) and disrespect for authorities (such as parents, the law, or one's country) is immoral, even liberals and libertarians have to admit that those judgments are moral in the sense that they are intended to be about morality. What liberals and libertarians deny is not that these judgments are moral judgments (as opposed to aesthetic, prudential, conventional, legal, or religious judgments). Instead, what they deny is only that these judgments are accurate, acceptable, or true. Otherwise, liberals and libertarians would not really be disagreeing with conservatives but only talking about a different, nonmoral topic.

Moreover, conservatives who claim harmless immorality see these judgments as moral in nature and basically similar to paradigm moral judgments about harm, injustice, and dishonesty. It might not be clear why they see their judgments as moral, but the fact that enough of them do intend their judgments to be about morality and do classify such judgments as moral despite knowing that they are not about harm is enough to show that harm cannot unify all judgments that are moral in the way defined above in section 14.1.

Finally, even if all moral judgments did somehow concern harm, that would not be enough to unify morality unless the notion of harm itself is unified. Harms can be physical (such as death, injury, and disease) or mental (such as loneliness and sadness). Harm is also sometimes defined to include "injustice and violation of [moral?] rights" (Haidt et al. 1993, 613), such as stealing fruit that its owner never notices or needs. Robinson and Kurzban (2007, 1866) add "intangible harms, such as the extent of intrusion of privacy." Gray and Wegner (2011, 258) extend harm so far that "even seemingly non-harm domains

are understood in the currency of harm." Gray, Young, and Waitz (2012, 16) even include "spiritual destruction," whatever that is. To cover all kinds of moral judgments, the term "harm" has to be extended so broadly that a happy and healthy secret masturbator is seen as harming himself (at least by those who judge that masturbation is immoral). This stipulation conflicts with common language and causes confusion. It also cannot show that the various harms are unified in any way that could generate any useful generalizations that are distinctive of morality.

That is why moral judgments cannot be unified by harm to others. Of course, moral judgments still might be unified by some other feature of content. However, it is hard to imagine what that other feature of content would be. In the absence of any plausible proposal, I conclude that moral judgments are not unified by their content.

14.3. Is Morality Unified by Its Neural Basis?

What else could unify moral judgments? Since salt comes in various colors but is unified by a shared chemical structure, one might hope to find a physical basis for all and only moral judgments. Such a common physical basis would presumably be located in the brain. Where else? Indeed, some studies of moral judgments might seem to suggest a shared neural basis that is distinctive of moral judgments. Many of Moll's early studies aimed to distinguish the neural basis for moral judgments in contrast with nonmoral judgments (Moll 2005, quoted above). However, recent neuroimaging evidence clearly tips the balance in favor of disunity—that is, in favor of the thesis that no neural system is both distinctive of moral judgments and also shared by all moral judgments. This section surveys a selection of these findings.

14.3.1. Ideals versus Requirements and Prohibitions

Different brain processes have been discovered for moral judgments about ideals in contrast with requirements and prohibitions. In Moll et al. (2006), subjects were given $128 and knew that they could leave with any of it that they did not give away. These subjects were asked to read about real charitable organizations. Some of these charities dealt with euthanasia, abortion, children's rights, the death penalty, gender equality, war, and nuclear power, so they raised controversial moral issues. Subjects were then given the chance to respond yes" or no to a choice of real payoffs to themselves and to a certain charity. For example,

Pure monetary reward: Yes to +$2 to you and $0 to the charity
Costly donation: Yes to −$2 from you and give +$5 to the charity
Noncostly donation: Yes to $0 to you and +$5 to the charity
Noncostly opposition: No to $0 to you and +$5 to the charity
Costly opposition: No to +$2 to you and +$5 to the charity

One might wonder why people would turn down $2 for themselves in order to prevent $5 going to a charity. The answer is moral judgment. Participants who see abortion as morally prohibited might accept some sacrifice in order to avoid enabling abortion. More generally, subjects who opposed donations to certain charities either at some cost or no gain to themselves presumably based that decision on a moral judgment that those charities violated some moral prohibition; and subjects who gave donations to certain charities at some cost to themselves presumably based that decision on a moral judgment that those charities further some moral ideal.

Moll et al. found that these judgments about moral prohibitions and moral ideals were associated with activation in distinct brain regions. Subjects who donated on the basis of moral ideals showed higher activation in the subgenual region. In contrast, subjects who opposed donations on the basis of moral prohibitions showed higher activation in the lateral orbital frontal cortex. Some regions were activated by both moral ideals and moral prohibitions, including the ventral tegmental area, the striatum, and the anterior prefrontal cortex. However, the ventral tegmental area and the striatum were also activated by pure monetary rewards as well as other kinds of reinforcement expectancy, so they were not activated *only* for moral judgments. The anterior prefrontal cortex was not activated by *all* moral judgments, because it was not activated by noncostly donations or noncostly opposition, so it seems to be related not to moral judgment but to cost to the chooser. Overall, no region of the brain is common and peculiar to both moral ideals and moral prohibitions. Those moral judgments seem to be made by different systems.

Critics might deny either that these donation choices were based on moral judgments or that we should ever have expected moral ideals to have the same neural basis as moral requirements or prohibitions. Even if moral ideals have a distinct neural basis, all moral requirements or prohibitions still might share a common and peculiar neural basis. That would not show that all moral judgments are unified in the brain, but it would still be important.

14.3.2. KINDS OF REQUIREMENTS AND PROHIBITIONS

Another study disconfirms that more limited hypothesis. Schaich Borg, Lieberman, and Kiehl (2008) asked participants in a brain scanner to memorize and recall statements that described neutral controls and target acts that involved three kinds of disgust: pathogen disgust, sexual disgust, and moral disgust. For example,

Control: You have dinner with your sister.
Pathogen disgust: You touch your sister's feces.
Sexual disgust: You give your sister an orgasm.
Moral disgust: You push your sister in front of a bus.

The assumption was that people who memorized these statements and repeated them would be making implicit moral judgments. The behavioral results are crucial here. Subjects rated the acts in sexual and moral disgust statements as equally immoral, but the acts in pathogen disgust statements were rated as much less immoral or not immoral at all. Schaich Borg et al. (2008) also found significant activations in many different brain areas. Pathogen disgust was related to activation in the amygdala, orbital frontal cortex, left inferior frontal gyrus, precuneus, visual cortex, and left fusiform gyrus. Sexual disgust was related to activation in the anterior cingulate, medial prefrontal cortex, insula, posterior cingulate, temporoparietal junction, and middle temporal gyrus. Moral disgust was related to activation in only the inferior parietal lobule. In addition, all disgust stimuli were related to activation in some common areas, including the medial prefrontal cortex, middle temporal gyrus, amygdala, and occipital gyrus.

The crucial point here is that no neural region is common and peculiar to all moral judgments—that is, to judgments regarding moral disgust and sexual disgust but not pathogen disgust. Some brain regions were related to *all* moral judgments in this study, but these areas were related to pathogen disgust where subjects did not make moral judgments. Other brain regions were related to *only* moral judgments—for example, the insula was related to only sexual disgust, about which subjects did make moral judgments—but those areas were not activated by other moral judgments, such as those related to moral disgust. Crucially, no brain regions were common and peculiar to all kinds of moral judgments. In order to show that morality is unified in an interesting way that enables distinctive and significant generalizations, some neural region would have to be both common and peculiar to all moral judgments. Hence, this study supports the disunity thesis.

Unfortunately, the study by Schaich Borg et al. (2008) did not ask subjects to make explicit moral judgments but only to memorize and recall statements that would presumably invoke implicit moral judgments. To see whether this makes a difference, we designed a study (Parkinson et al. 2011) that directly tests the unity hypothesis for explicit moral judgments.

The first step was to construct a set of stimuli that cleanly separates different areas of morality. Over the course of three pilot studies, we developed balanced scenarios that described actions that were (physically) harmful but neither dishonest nor disgusting, separate acts that were dishonest but neither harmful nor disgusting, and a third group of acts that invoked (sexual) disgust but were neither harmful nor dishonest. For example, the second group included undiscovered lies, and the third group included necrophilia without deception. We developed these stimuli so that they would be ambiguous in the specific sense that at least 30 percent of subjects found the act to be morally wrong and at least 30 percent found the act to be not morally wrong. This enabled us to compare brain activations when participants actually judged that an act was morally wrong to the brain activations when they judged that an act was not morally wrong. A final

group for comparison included neutral stimuli that were neither harmful nor dishonest nor disgusting and were not judged to be morally wrong.

Our findings confirmed our hypotheses. Harmful acts that were judged morally wrong compared to neutral scenarios were associated with increased activity in the left dorsal lateral prefrontal cortex, anterior cingulate cortex, supplementary motor area, inferior parietal lobe, superior temporal sulcus, and thalamus. In contrast, when participants judged dishonest acts to be morally wrong, there was bilateral activity in the dorsal medial prefrontal cortex (dmPFC), temporal parietal junction extending superiorly into the inferior parietal lobe, and posterior cingulate cortex as well as increased activity in the left dorsal lateral prefrontal cortex. When disgusting acts were judged to be morally wrong, we observed increased activity in the bilateral dorsolateral prefrontal cortex, dmPFC, amygdalae, anterior cingulate cortex, and posterior cingulate cortex, as well as the right temporal pole and left inferior frontal operculum/anterior insula. In summary, none of these areas were common and peculiar to all and only judgments of moral wrongness.

We did, however, find one location of activation that was common to all areas of moral judgment. A conjunction analysis revealed one twenty-eight-voxel cluster in the dmPFC that was independently activated in the comparison of each area of moral wrongdoing compared to neutral scenarios. To investigate whether this region was activated specifically by moral wrongdoing, we applied this cluster as a mask to each individual's data to extract the average response for the not-wrong scenarios in each category type. This secondary analysis revealed that dishonest and harmful scenarios judged to be not-wrong also activated this region more strongly than neutral scenarios. Thus, the activity observed in this region of the dmPFC seems to reflect a process engaged by the moral scenarios that is not peculiar to a moral judgment that the act is wrong. This interpretation is consistent with research showing that activity in this region of the dmPFC is modulated by ambiguity in both social and nonsocial contexts (van Overwalle 2009). By design, our moral scenarios were ambiguous in a way that our neutral scenarios were not, so ambiguity may explain the observed dmPFC activation across areas in the comparison between moral and neutral scenarios. Regardless of the exact mechanisms, the overlapping activation in the dmPFC seems to reflect some process or processes that are not peculiar to moral judgments that something is morally wrong.

Although our study focused on harm, dishonesty, and disgust, other studies have investigated the neural basis of moral judgments about fairness or justice. Robertson et al. (2007) found that areas of the brain reacted differently to sensitivity to harm versus sensitivity to unfairness and dishonesty. Hsu, Anen, and Quartz (2008) found that judgments of equity activated the insula, whereas judgments of efficiency activated more the putamen. These studies and others provide reason to expect that our results (Parkinson et al. 2011), as well as those of Schaich Borg, Lieberman, and Kiehl (2008), would generalize to other areas of morality.

Critics might respond that, even if not all moral judgments are unified by a brain mechanism, there still might be a unified brain mechanism underlying moral judgments of harmful acts in particular. That would not be enough to unify moral judgments in general, but it might satisfy some theorists who reject other kinds of judgments as not really moral judgments. However, Greene et al. (2001, 2004) found different brain activations for moral judgments of harmful acts in personal dilemmas as opposed to impersonal dilemmas. Schaich Borg et al. (2006) also found different brain activations for moral judgments of harmful acts when the harm was intended than when it was foreseen as an unintended side effect. Thus, not all moral judgments activate the same brain regions even when the moral judgments are all specifically about harm. This suggests that a prototypical neural response for a particular judgment, should it exist, will exist at a fine-grained level of analysis many subcategories below "moral judgment" (e.g., a specific kind of harm in a particular context from a distinct perspective). Accordingly, there appears to be slim hope of finding a prototypical neural correlate common and peculiar to all moral judgments.

Critics might reply that I require too much. The circulatory system seems unified, but it is distributed throughout the body instead of being located in only one part of the body. Thus, there can be unity without colocation. Similarly, even if moral judgments are not associated with activity in a single brain area, they still might have a distributed kind of unity. Nonetheless, it is one thing to say that the system might be unified and quite another to say that it actually is unified. If the circulatory system did not include physical connections or similar cell types, then it would not be unified at the physical level. In contrast, there is no evidence that moral judgments are related to anything in the brain that is remotely like the blood, veins, and arteries that make up the circulatory system. Hence, this analogy cannot show that moral judgments are unified at the physical level. Of course, the circulatory system is also unified at the functional level, because its various parts function to spread nutrients and other chemicals through the body. Similarly, even if moral judgments are not unified at the physical or neural level, they still might be unified at the level of function. That is a separate hypothesis that we need to consider next.

14.4. Is Morality Unified by Its Function?

The most popular proposal for unifying moral judgments appeals to function. As we saw above, Greene and Haidt agree with my conclusion in the preceding section when they write, "There is no specifically moral part of the brain" and "Morality is probably not a 'natural kind' in the brain" (2002, 522 and 523). Haidt (2011) also emphasizes that the content of morality covers several distinct "foundations." Nonetheless, Greene and Haidt still both claim that moral judgments share a distinctive function that unifies moral judgments at a different, nonneural level (Haidt 2011; and Greene 2013, 23;

discussed below). Thus, these moral theorists do not agree with my radical thesis that moral judgments are not unified at *any* level.

If moral judgments are unified by some function, then we can cite that function to distinguish moral judgments from nonmoral judgments and also to explain why so many people make moral judgments as well as why they make moral judgments in some cases but not in other cases. However, we still need to determine precisely which function can play these explanatory roles.

In order to specify the function of moral judgments, it seems natural to cite some widespread problem that moral judgments help to solve. Most proposed functional accounts of moral judgments specify the function in this way, but these accounts differ in the particular problem that moral judgments are supposed to solve. The challenge is then to find a single problem that all and only moral judgments solve, are intended to solve, or evolved to solve.

In order to fit the wide variety of moral judgments surveyed above in section 14.2, the relevant function, purpose, or problem must be characterized very broadly. One general problem for all societies is to control conduct by members of the group. This problem can be built into a definition of morality like this: "Morality is considered as the sets of customs and values that are embraced by a cultural group to guide social conduct" (Moll et al. 2005, 799). Although this sort of definition is popular, it cannot be adequate, because it is both too narrow and too broad.

First, it is too narrow, because some moral judgments are not about "social conduct." Many moralists, including Kant (1797, 82–107), condemn masturbation and suicide as immoral even in cases where they know that nobody else is involved or affected. Indeed, these and other moral duties to self have been a major topic in moral theory for centuries. The whole point of this debate is to determine whether acts can be immoral independent of other people. Whichever position you happen to agree with, the judgments on opposing sides are clearly moral (as opposed to aesthetic, economic, legal, etc.). However, the moral judgments that favor duties to self explicitly do not refer to "social conduct." These moral judgments are missed by any definition that limits morality to "social conduct."

Moll's definition is also too broad in other respects. Notice that this definition explicitly includes "customs." However, Turiel (1983) and his followers found that most people distinguish customs or conventions from morality. Many of the conventions or customs that Turiel's subjects distinguished from morality were based on values (even if not moral values) and were "embraced by a cultural group to guide social conduct." After all, "social conduct" includes everything that we do together with other people, and societies often value their own conventions or customs for various reasons.

One simple example is rules of dancing. These rules guide conduct that is social insofar as more than one person is dancing, and different societies value their own forms of dancing. Thus, rules of dancing do seem to be "customs and values that are embraced by a cultural group to guide social conduct," as the above definition says. Nonetheless, idiosyncratic dancing that breaks the social norms is often seen as ugly or weird but not

as immoral. Even dancers that hurt their partners by stepping on toes are judged to be uncoordinated but not immoral (unless they do so intentionally or carelessly). Some people might judge sexually explicit dancing to be immoral, but the relevant example here is unconventional dancing that is harmless and nonsexual but still judged to be ugly but not immoral. This example illustrates the general problem that there are many kinds of customs and many kinds of values. Dancing serves a kind of value (aesthetic) that is distinguished from morality. That is why we cannot define moral judgments simply as "the sets of customs and values that are embraced by a cultural group to guide social conduct."

Another instance of the same problem involves the rules of different languages. People who speak the French language might value placing adjectives after nouns, because that order seems more natural to them; and people who speak English might value placing adjectives before nouns, because that order seems more natural to them. These claims might seem silly, but the point here is only that many people value the customs of their own societies, but they still do not think that it is immoral for other societies to have other customs. It is wrong linguistically but not morally to place the adjective after the noun when speaking English.

Next compare laws that govern driving. Some people from Australia and the United Kingdom argue that it is better to drive on the left side of the road, because most people are right handed, so left-side driving allows more drivers to keep their dominant hand on the steering wheel when they shift gears in a manual transmission. Such judgments might seem silly, but that does not matter here. The point here is only that these people value their own customs, but they still do not think that it is immoral to drive on the right in the United States. Of course, it is immoral to intentionally drive on the right in the United Kingdom when this endangers other drivers, but that is a separate matter. People who value driving on the left do not think that it is immoral to drive on the right in countries where right-side driving is safe and legal. Indeed, they usually think that it is immoral to drive on the left in those countries. They see right-side driving as imprudent or inefficient as a custom but not as immoral when it is customary. Thus, this example and the others show that we cannot define moral judgments simply in terms of "the sets of customs and values that are embraced by a cultural group to guide social conduct." That definition includes too much.

In order to avoid these problems, some moral theorists introduce selfishness or self-interest as a target for the function to morality. Here is Greene (2013, 23): "Morality is a set of psychological adaptations that allow otherwise selfish individuals to reap the benefits of cooperation." Compare Haidt (2011, 270): "Moral systems are interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate self-interest and make cooperative societies possible." Thus, Greene and Haidt seem to agree that moral judgments evolved to serve a distinctive function of overcoming selfishness in order to enable cooperation.

This story sounds plausible in the abstract until you look more closely. To show why such definitions cannot really unify moral judgments, I will focus on Haidt's more complex version. Notice three features of his definition.

First, Haidt ascribes functions not to moral judgments but rather to "moral systems" as wholes. The problem is that moral systems include many parts, and the relations among those parts are exactly what is at issue here. Haidt might be claiming only that (A) any system of rules that can legitimately be called a moral system must have at least some parts that serve the specified functions. That claim (A) is compatible with recognizing that many other parts of the system do not serve these functions at all. But then it will not be true that (B) any rule that is part of a moral system must serve those functions. Claim (B) could be used to argue that a certain judgment is not a moral judgment, but claim (A) is useless for that purpose. For that reason, I will focus on claim (B) while recognizing that Haidt might have intended only claim (A).

Second, Haidt claims that moral systems "suppress or regulate self-interest and make cooperative societies possible." He does not explicitly say that nothing else serves this function. Hence, his official definition implies that (C) all moral judgments "suppress or regulate self-interest and make cooperative societies possible" but not that (D) *only* moral judgments "suppress or regulate self-interest and make cooperative societies possible." The problem, as before, is that claim (D) could be used to argue that a certain judgment is a moral judgment, but claim (C) is useless for that purpose. For that reason, I will focus on claim (D) while recognizing that Haidt might have intended only claim (C).

Third, Haidt's definition actually ascribes two functions instead of just one. To "suppress or regulate self-interest" is not the same as to "make cooperative societies possible." Which of these is supposed to be the function that unifies morality? This is unclear, so I will discuss them separately.

Consider first Haidt's claim that moral systems "make cooperative societies possible." What exactly does this mean? A minimal interpretation is simply that cooperative society would not be possible without at least some morality. Totally amoral people could not cooperate or live together for long. That seems plausible. Nonetheless, it cannot help us distinguish moral from nonmoral judgments. Even if total amorality would undermine society, there could still be many parts of morality that are not needed to "make cooperative societies possible." If some parts of morality are essential to cooperative society, but other parts are not, then this function is not common and peculiar to all moral judgments.

What would be needed to determine which judgments are moral is the stronger claim that all (and only) moral judgments "make cooperative societies possible." That stronger claim is, however, implausible. The problem can be illustrated by many permissive moral judgments. The judgment that I am morally permitted to spend my money on a yacht is moral in nature even though it is not needed to "make cooperative societies possible." A cooperative society could get along perfectly well, perhaps even better, without this moral judgment, even if it needs some rules in the general area of property.

In addition, many moral prohibitions are not really necessary for society. Just recall the survey of moral judgments in section 14.2, and it will be easy to find many moral judgments that are not necessary for cooperative society. For example, cooperative societies can survive without moral judgments against cannibalism, either because they practice limited cannibalism or because they have an aversion to cannibalism that prevents the practice without leading to a moral judgment.

Another problem is that several kinds of nonmoral rules or judgments are also needed to make cooperative society possible. We saw examples above: language and law. Without any rules of language, we could not communicate in ways that are needed for us to form anything like the societies we have; but it is still not immoral to break those rules of language (even intentionally). Similarly, without any laws, we could not live together in cooperative societies; but we can still judge an act to be illegal without judging it to be immoral. Admittedly, many individual rules of language and law are not necessary to "make cooperative societies possible." It is only language and law as a whole that are needed to "make cooperative societies possible." But that is precisely the situation with morality. Hence, this function cannot be used to show what is distinctive about morality.

The other function that Haidt ascribes to morality is to "suppress or regulate self-interest." (Notice that Greene also alludes to "selfish individuals.") However, some moral judgments seem not to function to "suppress or regulate self-interest"—at least not on an individual level. After all, the judgment that I am morally permitted to spend my money on a yacht is moral in nature even though it does not "suppress or regulate self-interest." Haidt could respond that his definition was meant to apply not to such permissive moral judgments but only to moral judgments about what is forbidden or obligatory. However, this admits that his definition does not work for all moral judgments, since permissive moral judgments constitute a large group of moral judgments.

In addition, many moral judgments about obligations also do not "suppress or regulate self-interest." If I promise my wife to go for a walk with her, then I might judge that my promise creates a moral obligation to go for a walk with her, even while I very much want to do so and it is in my self-interest to do so. On many occasions like this, morality and self-interest coincide. Admittedly, the general rule that we ought to keep our promises does conflict with self-interest in some situations, but such occasional conflicts cannot help us distinguish moral judgments from nonmoral judgments in general.

Finally, it is not *only* moral judgments that "suppress or regulate self-interest." Rules of games also have that function, since it would often be in my self-interest to move my rook on a diagonal when I am playing chess. The same goes for language. I might want to use words contrary to the rules of grammar, but doing so is still ungrammatical, so grammar can also "suppress or regulate self-interest." (Lewis Carroll made this point with Humpty Dumpty.) Indeed, it might be a general feature of all rules (whether moral or not) that they sometimes (though not always) regulate self-interest. Nonetheless, this function of regulating self-interest does not turn rules of chess or grammar into moral

rules. We judge chess moves to be illegal and utterances to be ungrammatical without judging them to be immoral. Thus, even when moral judgments do "suppress or regulate self-interest," that function cannot be used to define moral judgments.

It might seem that I cheated by separating the two functions that Haidt mentioned together, so let's consider conjunctions and disjunctions. Can moral judgments be identified as judgments that serve both to regulate self-interest and also to make cooperative society possible? No, because I gave examples of moral judgments that lacked one of these functions, so those judgments also fail the conjunctive requirement. Next, can moral judgments be identified as judgments that serve either to regulate self-interest or to make cooperative society possible? No, because I gave examples of nonmoral judgments that have one of these functions, so those judgments also meet the disjunctive requirement. Thus, moral judgments cannot be unified by taking these two functions together.

To avoid these problems, moral theorists might refer not just to selfishness but also to other kinds of partiality. Warnock (1971), for example, argued that the purpose of morality and of moral judgments is to reduce liability to harm that arises from limited sympathy. The purpose could not be simply to reduce liability to harm in general, since hygienic rules about brushing your teeth reduce liability to harm but are not moral rules. The particular concern of morality, according to Warnock, was that we care about ourselves and our family and friends much more than we care about strangers. This limited sympathy then leads us to do acts that harm others, so we are all in danger when lots of other people have such limits on their sympathy. Morality is supposed to solve this problem by having rules that apply to all people equally. One should not lie, break promises, kill, or rape anyone—even strangers for whom one has no sympathy.

This picture might seem plausible in many cases, but it does not cover all of the judgments that people usually classify as moral. Consider rules about purity. A judgment that it is morally wrong to commit homosexual sodomy will reduce sympathy for homosexuals and the difficulties that they have when prevented from freely expressing their love. Batson and his collaborators (1999) found, for example, that people who judge homosexuality to be immoral do not help out needy homosexuals, even when they just need money to visit their grandparents. Similar points apply to moral judgments about hierarchies and loyalty. When people judge that we must obey our parents and bosses or that we have special duties to our in-group, this judgment will presumably reduce sympathy for people who are not in our in-group or whom our parents or bosses order us to hurt, neglect, or despise. Moral judgments about hierarchies and loyalty function not to reduce limits on sympathy but, instead, to impose limits on sympathy for underclasses and outsiders.

Next consider retributive justice. The rule that punishers must take an eye for an eye and a tooth for a tooth seems to undermine sympathy for people who have committed crimes. On this rule, officials should punish them in proportion to their crime regardless to how much sympathy anyone feels for the criminal. Even those who reject

retributivism must recognize such retributivist judgments as moral judgments instead of aesthetic, prudential, legal, or conventional judgments. That shows that not all moral judgments function to expand sympathy. More generally, the problem in morality sometimes is too much sympathy rather than too little. The classic expression that hard cases make bad law does not, as many people assume, mean that difficult cases make bad law. Instead, it refers to cases where you need to be hard on someone. For example, if you need to punish or civilly penalize someone who has done something wrong, but you feel really sorry for her (perhaps because she is sick or had a child die), then you might bend the law so that you will not have to punish her or make her pay damages in a civil case. Here sympathy gets in the way of justice. Since justice is an area of morality, morality does not always serve the function of overcoming limits on sympathy, as Warnock claimed.

Let's consider one more recent attempt to unify morality by its function. Robinson and Kurzban (2007, 1865–1866) claim that "what makes an act immoral is the concurrent belief that those who perform the act should be punished." This proposal is not entirely clear, but it can be interpreted as suggesting that the function of judging an act to be immoral is to distinguish agents who should be punished from those who should not and perhaps also to justify punishing the certain agents. Nonmoral judgments (such as aesthetic, prudential, and conventional judgments) are not supposed to justify punishment in the same way. If so, we can identify moral judgments as those that are intended to tell us who "should be punished."

Unfortunately, this account is again both too broad and too narrow. It is too broad because, for example, there can be fines for illegal parking and penalties for failing to file tax forms on time, even when those acts are not judged to be immoral. In the classic terminology, mala prohibita are punishable, even though they are not mala in se or immoral in themselves. Moreover, even in cases of mala in se, such as murder, there might be strong reasons for criminal laws against some actions, such as active euthanasia, even in circumstances where those acts are not immoral, simply because allowing such acts would lead to too many mistakes. Conversely, Robinson and Kurzban's definition is also too narrow because many acts are seen as immoral even if nobody should punish the agent. Imagine that your neighbor's wife lies to him, and he immediately forgives her. Who should punish her? Him? But he forgave her. You? But it's none of your business. It seems at least possible that, even if her lie was immoral, there is nobody who should punish her, as Robinson and Kurzban's definition requires. (Perhaps she was liable to punishment by her husband before he forgave her, but their definition requires that someone should punish her, and even he should not punish her if he should forgive her as he did.) In addition, recall permissive moral judgments, such as the judgment that it is not immoral to buy a yacht. Permissive judgments do not justify punishment. That is what makes them permissive. Remember also judgments of moral ideals, such as that it is nice to give the tenure gift described above. The judgment that the gift was nice, good, ideal, or supererogatory can be intended as a moral judgment as opposed to a religious, aesthetic, economic, conventional, or legal judgment, even if no punishment would have

been justified if the gift had not been given. Thus, neither all nor only moral judgments justify punishment.

Kurzban admits as much in a later development of his views. DeScioli and Kurzban (2009, 282) "argue for two distinct components of moral cognition: One subsystem regulates one's own behavior (conscience) and another mechanism is specialized for judging others (condemnation)." Of course, if different parts of morality have two distinct functions, as they claim, then morality is no more unified than jade is unified when it has two subtypes, jadeite and nephrite. Maybe the claim is instead that each moral judgment has both functions: conscience and condemnation. However, condemnation simply generalizes punishment, so it still cannot account for permissive judgments or ideals, discussed above. Consider also heroism. We can judge someone a moral hero, and that is a moral judgment, even if we do not condemn people who do not live up to such high standards. This case and many others like it show that condemnation plays a role in only some but not all moral judgments. Moreover, we condemn people for nonmoral wrongs, such as ungrammatical sentences, stupidity, invalid arguments, and illegal parking. A special kind of condemnation might seem appropriate for violations of moral prohibitions, but then the problem is to specify that special kind of condemnation (without circularly calling it "moral condemnation"). The general notion of condemnation by itself is not enough to enable us to identify moral judgments, because we condemn lots of things other than immorality. Adding the notion of conscience cannot help here because conscience is defined so broadly as that which "regulates one's own behavior." We can use hygienic rules, such as "Brush your teeth," to regulate our own behavior without judging it immoral to skip brushing our teeth because we are too tired. The same goes for rules of grammar, games, logic, and prudence. In all such cases, we can regulate behavior and condemn rule violations without judging anything immoral. Thus, this new proposal by DeScioli and Kurzban also cannot capture what is distinctive about morality.

Many more possible proposals about the function of morality deserve consideration here. However, instead of extending the list, I want to step back and raise the general question of why functional definitions of morality all fail. Evolutionary psychologists cite evidence that different moral judgments arose from different evolutionary pressures at different times. For example, Lieberman and Hatfield (2006) argue that pathogen disgust (connected to prohibitions on cannibalism) arose from one type of evolutionary pressure; then sexual disgust (connected to prohibitions on incest) arose from a different kind of evolutionary pressure; and then "moral" disgust (such as finding nonsexual sadism disgusting) arose from yet another evolutionary pressure. Moreover, it is hard to see how moral rules against lying or promise breaking could arise before language or why moral rules against harm would not have arisen in some form before language, since our ancestors were subject to harmful aggression long before they could talk. In these cases and others, then, different moral judgments probably arose from different evolutionary pressures in different circumstances. If that is correct, then it should come as no surprise that different kinds of moral judgments serve different evolutionary functions as well as

different contemporary functions. This evidence thus provides reason to doubt that any single function could unify all areas of moral judgment. Morality is just too diverse in its history.

Admittedly, there is a large body of fascinating and plausible work about the evolution of morality that I have not discussed. However, most of that literature does not even try to show that all and only moral judgments serve a common and distinctive function. Instead, most studies focus on one or a few particular kinds of moral judgments, as is proper (see section 14.5). Hence, this research cannot undermine my conclusion that different moral judgments have different functions. Indeed, surveying the larger literature would further support the conclusion that morality has no single unifying function, as Machery and Mallon (2010) have argued.

14.5. Does Disunity Matter?

My overall conclusion is that morality is not unified. This conclusion can meet with either of the two prototypical philosophical replies: (1) Oh yeah? (2) So what?

Some critics respond that I have not proven this conclusion. Granted. Since I have not and cannot survey all possible proposals for unifying morality, my conclusion is best characterized as raising a question: If there is something else that unifies all and only moral judgments, what is it? Opponents need to put forward a precise proposal. Only then can we assess it.

Another common response is this: Fine, but who cares? All scientists and philosophers should care. If morality really is like memory, as I suggested, then scientists and philosophers can make more progress if they stop trying to study morality all at once. Here's why.

In many experiments, researchers lump together moral judgments and seek contrasts between moral and nonmoral judgments. These experimenters do not draw distinctions among different kinds of moral judgments or test these kinds separately. Instead, they throw them all into a single bin. I suggest, instead, that scientists should isolate smaller classes of judgments within one region of the map, carefully distinguish from other types of judgments not on the basis as whether they count as moral or not, but rather on the basis of their content and context. Then scenarios should be presented from a single perspective either in the first person, second person, or third person consistently. And then scientists should look for the neural basis or the evolutionary origin or the psychological process behind that smaller class of judgments.

After a series of such smaller-scale experiments, of course, we can ask whether this basis or origin in each area extends to all or only moral judgments. We might eventually find some unifying feature. I doubt it, but I cannot rule out the possibility on the basis of what I have argued here.

This bottom-up method has several advantages. It prevents both false positives and false negatives. First, if researchers lump together different types of moral judgments in

empirical studies, they might miss a lot of interesting conclusions. For example, if one class of moral judgments in the mix activates an area of the brain more highly than the baseline, but then other moral judgments activate that same area of the brain less than the baseline, then researchers will miss both of these significant brain activation levels, because they will cancel out each other. Similarly, we might find no significant difference between moral and nonmoral conditions, even though a subset of moral judgments is related to a significant activation increase, simply because the increased activation from a subset of our stimuli is not statistically significant when thrown together with all the other types of moral judgments.

The bottom-up method also helps to avoid false positives. For example, if researchers present ten moral scenarios each from five different areas, and they find an increased activation for that total set of compared to a neutral baseline, then they might conclude that all of the moral judgments tend to activate those areas. However, that entire result might come merely from one or two classes of moral judgments out of the five kinds, even though the other moral judgments do not activate these brain areas at all. If they assume that those other judgments will activate that area, then they will fall into the trap of false positives.

To avoid such mistakes and to make more progress, moral science needs to shift towards taxonomic rigor. The first step is to accept that moral judgment is not unified. I hope that this chapter has convinced readers of that fact. The next step is for researchers to begin to create a working taxonomy that informs future paradigm design and data interpretation. Any such taxonomy should be based on detailed empirical research combined with careful philosophical analysis. By committing to a taxonomic approach, the field can progress toward a deeper understanding of the component processes that define the myriad judgments that we call "moral."

Note

1. Sinnott-Armstrong (2008) and Sinnott-Armstrong and Wheatley (2012, 2013) argue against other ways to unify morality. In this chapter, I draw heavily on that previous work. I am deeply indebted to Thalia Wheatley for our joint research on these issues. I am also grateful to many other friends for comments on this series of articles and to Matthew Liao and Peter DeScioli for comments on an earlier draft of this chapter.

References

Batson, C. D., R. B. Floyd, J. M. Meyer, and A. L. Winner. 1999. "And Who is My Neighbor? Intrinsic Religion as a Source of Universal Compassion." *Journal for the Scientific Study of Religion* 38: 445–457.

DeScioli, P., and R. Kurzban. 2009. "Mysteries of Morality." *Cognition* 112: 281–299.

Flanagan, O. 1991. *Varieties of Moral Personality*. Cambridge, MA: Harvard University Press.

Gert, B. 2005. *Morality: Its Nature and Justification*. New York: Oxford University Press.

Gray, K., and D. M. Wegner. 2011. "Dimensions of Moral Emotions." *Emotion Review* 3: 258–260.

Gray, K., L. Young, and A. Waitz. 2012. "Mind Perception Is the Essence of Morality." *Psychological Inquiry* 23: 101–124.

Greene, J. 2013. *Moral Tribes: Emotion, Reason, and The Gap between Us and Them*. New York: Penguin.

Greene, J., and J. Haidt. 2002. "How (and Where) Does Moral Judgment Work?" *Trends in Cognitive Science* 6: 517–523.

Greene, J. D., L. E. Nystrom, A. D. Engell, J. M. Darley, and J. D. Cohen. 2004. "The Neural Bases of Cognitive Conflict and Control in Moral Judgment." *Neuron* 44: 389–400.

Greene, J. D., R. B. Sommerville, L. E. Nystrom, J. Darley, and J. D. Cohen. 2001. "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science* 293: 2105–2108.

Haidt, J. 2011. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York: Pantheon.

Haidt, J., and J. Graham. 2007. "When Morality Opposes Justice: Conservatives Have Moral Intuitions That Liberals May Not Recognize." *Social Justice Research* 20: 98–116.

Haidt, J., S. H. Koller, and M. G. Dias. 1993. "Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?" *Journal of Personality and Social Psychology* 65: 613–628.

Hsu, M., C. Anen, and S. Quartz. 2008. "The Right and the Good: Distributive Justice and Neural Encoding of Equity and Efficiency." *Science* 320: 1092–1095.

Kant, I. [1785] 1959. *Foundations of the Metaphysics of Morals*. Translated by L. W. Beck. Indianapolis: Bobbs-Merrill.

Lieberman, D., and E. Hatfield. 2006. "Lust and Disgust: Cross-Cultural and Evolutionary Perspectives." In *The Psychology of Love*, 2nd ed., edited by R. J. Sternberg and K. Weis, 274–297. New Haven: Yale University Press.

Machery, E., and R. Mallon. 2010. "Evolution of Morality." In *The Oxford Handbook of Moral Psychology*, edited by J. Doris and the Moral Psychology Research Group, 3–46. New York: Oxford University Press.

Mill, J. S. [1859] 2011. *On Liberty*. New York: Simon and Brown.

Moll, J., R. Zahn, R. de Oliveira-Souza, F. Krueger, and J. Grafman. 2005. "The Neural Basis of Human Moral Cognition." *Nature Reviews Neuroscience* 6: 799–809.

Moll, J., F. Krueger, R. Pardini, M. Zahn, R. de Oliveira-Souza, and J. Grafman. 2006. "Human Fronto-Mesolimbic Networks Guide Decisions about Charitable Donation." *Proceedings of the National Academy of Sciences* 103, no. 42: 15623–15628.

Parkinson, C., W. Sinnott-Armstrong, P. Koralus, A. Mendelovici, McV. Geer, and T. Wheatley. 2011. "Is Morality Unified? Evidence That Distinct Neural Systems Underlie Moral Judgments of Harm, Dishonesty, and Disgust." *Journal of Cognitive Neuroscience* 10: 3162–3180.

Robertson, D., J. Snarey, O. Ousley, K. Harenski, F. Bowman, R. Gilkey, and C. Kilts. 2007. "The Neural Processing of Moral Sensitivity to Issues of Justice and Care." *Neuropsychologia* 45: 755–766.

Robinson, P. H., and R. Kurzban. 2007. "Concordance and Conflict in Intuitions of Justice." *Minnesota Law Review* 91: 1829–1907.

Ross, W. D. 1930. *The Right and the Good*. Oxford: Clarendon Press.

Schaich Borg, J., D. Lieberman, and K. A. Kiehl. 2008. "Infection, Incest and Iniquity: Investigating the Neural Correlates of Disgust and Morality." *Journal of Cognitive Neuroscience* 20: 1529–1546.

Shweder, R. A., N. C. Much, M. M. Mahapatra, and L. Park. 1997. "The 'Big Three' of Morality (Autonomy, Community, and Divinity) and the 'Big Three' Explanations of Suffering." In *Morality and Health*, edited by A. M. Brandt and P. Rozin, 119–169. New York: Routledge.

Sinnott-Armstrong, W. 2008. "Is Moral Phenomenology Unified?" *Phenomenology and Cognitive Sciences* 7: 85–97.

Sinnott-Armstrong, W., and T. Wheatley. 2012. "The Disunity of Morality and Why It Matters to Philosophy." *Monist* 95: 355–377.

Sinnott-Armstrong, W., and T. Wheatley. 2013. "Are Moral Judgments Unified?" *Philosophical Psychology*. doi:10.1080/09515089.2012.736075.

Stich, S. 2006. "Is Morality an Elegant Machine or a Kluge?" *Journal of Cognition and Culture* 6: 181–189.

Sundby, A. 2010. "Study: Magnetic Waves Alter Moral Compass." CBS News, March 30. http://www.cbsnews.com/stories/2010/03/30/tech/main6347079.shtml.

Tulving, E. 2000. "Concepts of Memory." In *The Oxford Handbook of Memory*, edited by E. Tulving and F. I. M. Craik, 33–44. New York: Oxford University Press.

Turiel, E. 1983. *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.

Van Overwalle, F. 2009. "Social Cognition and the Brain: A Meta-analysis." *Human Brain Mapping* 30: 585–594.

Warnock, G. J. 1971. *The Object of Morality*. London: Methuen.