Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 3

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

Sample - Chapter 1 from Regard for Reason in the Moral Mind by Joshua May (Oxford University Press, 2018)

1

Empirical Pessimism

Reason is wholly inactive, and can never be the source of so active a principle as conscience, or a sense of morals.

– David Hume

1.1 Introduction

Moral evaluation permeates human life. We readily praise moral saints, admonish those who violate ethical norms, and teach children to develop virtues. We appeal to moral reasons to guide our own choices, to structure social institutions, and even to defend atrocities. But is this a fundamentally rational enterprise? Can we even rely on our basic modes of moral thought and motivation to know right from wrong and to act virtuously?

Empirical research may seem to warrant doubt. Many philosophers and scientists argue that our moral minds are grounded primarily in mere feelings, not rational principles. Emotions, such as disgust, appear to play a significant role in our propensities toward racism, sexism, homophobia, and other discriminatory actions and attitudes. Scientists have been increasingly suggesting that much, if not all, of our ordinary moral thinking is different only in degree, not in kind. Even rather reflective people are fundamentally driven by emotional reactions, using reasoning only to concoct illusory justifications after the fact. As Jonathan Haidt has put it, “the emotions are in fact in charge of the temple of morality” while “moral reasoning is really just a servant masquerading as the high priest” (2003: 852).

On such influential pictures, ordinary moral thinking seems far from a reasoned pursuit of truth. Even if some ordinary moral judgments are rational and reliable, brain-imaging research suggests that the intuitive moral judgments that align with commonsense morality are driven largely by inflexible emotional alarms instilled in us long ago by natural selection. The same apparently goes for our thinking about even the most pressing of contemporary moral issues, such as abortion, animal rights, torture, poverty, and climate change. Indeed, some

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 4
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

theorists go so far as to say that we can't possibly acquire moral knowledge, or even justified belief, because our brains have been shaped by evolutionary forces that can't track supposed "moral facts."

As a result, virtue seems out of reach because most of us don't know right from wrong. And it gets worse. Even if commonsense moral judgment is on the right track, distinctively moral motivation may be impossible or exceedingly rare. When motivated to do what's right, we often seem driven ultimately by self-interest or non-rational passions, not our moral beliefs. If our moral convictions do motivate, they are corrupted by self-interested rationalization or motivated reasoning. Scientific evidence suggests that people frequently lie and cheat to benefit themselves whenever they believe they can get away with it. Sure, we can feel empathy for others, but mostly for our friends and family. Those suffering far away don't stir our sentiments and thus don't motive much concern. When we do behave well, it's often to gain some reward, such as praise, or to avoid punishment. Doing what's right for the right reasons seems like a psychological rarity at best.

While theorists disagree over the details, there has certainly been an increase in scientifically motivated pessimism (a term I borrow from D'Arms & Jacobson 2014). These pessimists contend that ordinary moral thought and action are ultimately driven by non-rational processes. Of course, not all empirically informed philosophers and scientists would describe themselves as "pessimists." They may view themselves as just being realistic and view the optimist as a Panglossian Pollyana. But we'll see that the label of "pessimism" does seem apt for the growing attempts to debunk ordinary moral psychology or to pull back the curtain and reveal an unsophisticated patchwork in need of serious repair.

This book aims to defend a more optimistic view of our moral minds in light of our best science. Knowing right from wrong, and acting accordingly, is indeed difficult for many of us. But we struggle not because our basic moral beliefs are hopelessly unjustified—debunked by evolutionary pressures or powerful emotions—or because deep down we are all motivated by self-interest or are slaves to ultimately non-rational passions. Science can certainly change our conception of humanity and cause us to confront our biological and cultural limitations. Not all of commonsense morality can survive, but we should neither oversell the science nor commit ordinary moral thinking to the flames.

Ultimately, I argue for an *optimistic rationalism*. Ordinary moral thought and action are driven by a regard for "reason"—for reasons, reasonableness, or justifiability. Pessimists commonly point to our tendencies toward irrationality, but perhaps paradoxically it is often our irrationalities that reveal our deep regard for reason. If ordinary moral cognition had little to do with reason, then we

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 5

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

would not so often rationalize or provide self-deceived justifications for bad behavior. Driven by this concern to act in ways we can justify to ourselves and to others, moral knowledge and virtue are possible, despite being heavily influenced by unconscious processes and despite being sensitive to more than an action's consequences.

In this chapter, I'll introduce some key sources of pessimism about two core aspects of moral psychology. Some theorists are doubtful about the role of reason in ordinary moral *cognition* and its ability to rise to knowledge. Others are doubtful about the role of reason in moral *motivation* and our ability to act from virtuous motivation. After surveying a diverse range of opponents, I'll explain the plan in the coming chapters for defending a cautious optimism about our moral minds, and one that lies within the rationalist tradition.

1.2 Pessimism about Moral Cognition

1.2.1 Sources of pessimism

Contemporary moral philosophers have rightly turned their attention to the sciences of the mind in order to address theoretical and foundational questions about ethics. What is going through our minds when we condemn others or are motivated to do what's right? Is moral thinking a fundamentally inferential process or are sentiments essential? To test proposed answers to such questions, some philosophers are now even running their own experiments.

Unfortunately, though, philosophers and scientists alike have tended to hastily take this empirically informed movement to embarrass ordinary moral thinking or the role of reason in it. Ethical theories in the tradition of Immanuel Kant, in particular, have taken a serious beating, largely for their reverence for reason.

To be fair, Kantians do claim that we can arrive at moral judgments by pure reason alone, absent any sentiments or feelings. Contemporary Kantians likewise ground morality in rational requirements, not sentiments like resentment or compassion. Thomas Nagel, for example, writes: "The altruism which in my view underlies ethics is not to be confused with generalized affection for the human race. It is not a feeling" (1970/1978: 3). Instead, Kantians typically ground morality in reflective deliberation about what to do (Wallace 2006) or in reflective endorsement of one's desires and inclinations. Michael Smith, for example, argues that moral approbation expresses a belief about "what we would desire ourselves to do if we were fully rational" (1994: 185). Similarly, Christine Korsgaard writes that "the human mind . . . is essentially reflective" (1996/2008: 92), and this self-consciousness is required for moral knowledge and

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 6
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

virtue, for it allows us to make reasoned choices that construct our own identities. Morality, according to Korsgaard (2009), arises out of "the human project of self-constitution" (4), which involves a "struggle for psychic unity" (7).

Many empirical pessimists contend that reflection and deliberation do not play such a substantial role in our moral minds. Haidt even speaks of a "rationalist delusion" (2012: 103), and it's not difficult to see why. The study of moral development in psychology was dominated in the twentieth century by Lawrence Kohlberg (1973), who was heavily inspired by Kant. However, that tradition has largely fallen out of favor to make room for psychological theories in which emotion plays a starring role. Many psychologists and neuroscientists now believe that a surprising portion of our mental lives is driven by unconscious processes, many of which are automatic, emotional, and patently irrational or non-rational. Reasoning comes in to justify that which one's passions have already led one to accept. As Haidt has put it, "moral reasoning does not cause moral judgment; rather, moral reasoning is usually a post-hoc construction, generated after a judgment has been reached" (2001: 814).

This is the challenge from a brand of *sentimentalism* which contends that moral cognition is fundamentally driven by emotion, passion, or sentiment that is distinct from reason (e.g., Nichols 2004; Prinz 2007). Many now take the science to vindicate sentimentalism and Hume's famous derogation of reason. Frans de Waal, for example, urges us to "anchor morality in the so-called sentiments, a view that fits well with evolutionary theory, modern neuroscience, and the behavior of our primate relatives" (2009: 9). Even if reasoning plays some role in ordinary moral judgment, the idea is that sentiment runs the show (Haidt 2012: 77; Prinz 2016: 65).

Other critics allow that ordinary moral judgment can be driven by reason, but they attempt to *debunk* all or large portions of commonsense morality, yielding full or partial *skepticism*. Evolutionary debunkers argue that Darwinian pressures prevent our minds from tracking moral truths. Even if blind evolutionary forces get us to latch onto moral facts, this is an accident that doesn't amount to truly knowing right from wrong. As Richard Joyce puts it, "knowledge of the genealogy of morals (in combination with some philosophizing) should undermine our confidence in our moral judgments" (2006: 223; see also Ruse 1986; Rosenberg 2011).

Other debunkers align good moral reasoning with highly counter-intuitive intuitions consistent with utilitarian (or other consequentialist) ethical theories. Peter Singer (2005) and Joshua Greene (2013), for example, argue that moral thinking is divided into two systems—one is generally trustworthy, but the other dominates and should be regarded with suspicion. The commonsense moral

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 7
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

intuitions supporting non-utilitarian ethics can be debunked since they arise from unreliable cognitive machinery. Greene writes that our "anti-utilitarian intuitions seem to be sensitive to morally irrelevant things, such as the distinction between pushing with one's hands and hitting a switch" (328). These pessimists are utilitarian debunkers who argue that the core elements of ordinary moral judgment should be rejected, largely because they are driven by automatic emotional heuristics that place moral value on more than the consequences of an action. While some moral judgments are rational, and can yield knowledge or at least justified belief, most of our ordinary intuitions are not among them. Such utilitarians are often content with imputing widespread moral ignorance to the general population, which likewise renders virtuous action exceedingly rare.

Many debunkers conceive of moral cognition as facing a dilemma in light of the science. As Singer has put it:

> We can take the view that our moral intuitions and judgments are and always will be emotionally based intuitive responses, and reason can do no more than build the best possible case for a decision already made on nonrational grounds. [. . .] Alternatively, we might attempt the ambitious task of separating those moral judgments that we owe to our evolutionary and cultural history, from those that have a rational basis. (2005: 351)

It seems we can avoid wholesale sentimentalism only by undermining large swaths of ordinary moral thinking.

Whether by embracing sentimentalism or debunking, a pessimistic picture of ordinary moral thinking seems to result. The worry is that, if our best science suggests that our moral minds are driven largely by non-rational passions, then that way of thinking may be indefensible or in need of serious revision or repair. Now, sentimentalists frequently deny that their view implies that our moral beliefs are somehow deficient (see e.g., Kauppinen 2013; D'Arms and Jacobson 2014), and of course emotions aren't necessarily illicit influences. However, sentimentalists do maintain that genuinely moral cognition ultimately requires having certain feelings, which suggests that it's fundamentally an arational enterprise in which reason is a slave to the passions.

At any rate, I aim to provide a defense of ordinary moral cognition that allows reason to play a foundational role. First, I'll argue for an empirically informed *rationalism*: moral judgment is fundamentally an inferential enterprise that is not ultimately dependent on non-rational emotions, sentiments, or passions. Second, I'll advance a form of *anti-skepticism* against the debunkers: there are no empirical grounds for debunking core elements of ordinary moral judgment, including our tendency to place moral significance on more than an action's consequences.

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 8
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

1.2.2 Reason vs. emotion?

Philosophers and scientists increasingly worry that the reason/emotion dichotomy is dubious or at least fruitless. We of course shouldn't believe that reason is good and reliable while emotion is bad and biasing (Jones 2006; Berker 2009). Moreover, as we further understand the human brain, we find great overlap between areas associated with reasoning and emotional processing with apparently few differences. Like paradigm emotional processing, reasoning can be rapid and relatively inaccessible to consciousness. And emotions, like paradigm reasoning, aid both conscious and unconscious inference, as they provide us with relevant information (Dutton & Aron 1974; Schwarz & Clore 1983), often through gut feelings about which of our many options to take (Damasio 1994).

The position developed in this book is likewise skeptical of the reason/emotion dichotomy, but this won't fully emerge until the end. For now, let's begin by attempting to articulate a working contrast between reason and emotion.

Reasoning is, roughly, a kind of inference in which beliefs or similar propositional attitudes are formed on the basis of pre-existing ones. For example, suppose Jerry believes that Elaine will move into the apartment upstairs only if she has $5,000, and he recently learned that she doesn't have that kind of money to spare. Jerry then engages in reasoning when, on the basis of these two other beliefs, he comes to believe that Elaine won't move into the apartment upstairs. It's notoriously difficult to adequately characterize this notion of forming a belief "on the basis" of other beliefs in the sense relevant to inference (see, e.g., Boghossian 2012). But such issues needn't detain us here.

Some philosophers and psychologists define reasoning more narrowly as *conscious* inference (e.g., Haidt 2001: 818; Mercier & Sperber 2011: 57; Greene 2013: 136). This may capture one ordinary sense of the term "reasoning." The archetype of reasoning is indeed deliberate, relatively slow, and drawn out in a step-wise fashion. For example, you calculate your portion of the bill, weight the pros and cons of divorce, or deliberate about where to eat for lunch.

But there's no need to be overly restrictive. As Gilbert Harman, Kelby Mason, and Walter Sinnott-Armstrong point out: "Where philosophers tend to suppose that reasoning is a conscious process . . . most psychological studies of reasoning treat it as a largely unconscious process" (2010: 241). Moreover, ordinary usage and dictionary definitions don't make conscious awareness essential to reasoning, presumably because rule-governed transitions between beliefs can be a rather automatic, unconscious, implicit, and unreflective process. For example:

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 9
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

- You just find yourself concluding that your son is on drugs.
- You automatically infer from your boss's subtly unusual demeanor that she's about to fire you.
- You suddenly realize in the shower the solution to a long-standing problem.

These beliefs seem to pop into one's head, but they aren't born of mere feelings or non-inferential associations. There is plausibly inference on the basis of representations that function as providing reasons for a new belief. Reasoning occurs; it's just largely outside of awareness and more rapid than conscious deliberation.

Indeed, it is now common in moral psychology to distinguish conscious from unconscious reasoning or inference (e.g., Cushman, Young, & Greene 2010; Harman et al. 2010). The idea is sometimes emphasized by rationalists (e.g., Mikhail 2011), but even sentimentalists allow for unconscious reasoning, particularly in light of research on unconscious probabilistic inference (Nichols, Kumar, & Lopez 2016; see also Zimmerman 2013).

No doubt some of one's beliefs are formed without engaging in reasoning, conscious or not. Basic perceptual beliefs are perhaps a good example. You believe that the door opening in front of you retains a rectangular shape, but arguably you don't form this judgment on the basis of even tacit beliefs about angles in your field of vision. Rather, your visual system generates such perceptual constancies by carrying out computational work among mental states that are relatively inaccessible to introspection and isolated from other patterns of belief-formation (such states are often called *sub-personal*, although *sub-doxastic* [Stich 1978] is probably more apt [Drayson 2012]). As the visual experience of a rectangular door is generated, you believe that the door is rectangular by simply taking your visual experience at face value. So perhaps it's inappropriate to posit unconscious reasoning (about angles and the like) at least because the relevant transitions aren't among beliefs—not even tacit ones.

Nevertheless, some inferential transitions between genuine beliefs are unconscious. Within the category of unconscious mental processes, some generate beliefs on the basis of prior beliefs (e.g., inferring that your son is on drugs). Other belief-generating processes don't amount to reasoning or inference (e.g., believing that an opening door is constantly rectangular), at least because they are "subpersonal" or "subdoxastic."

What about emotion? There is unfortunately even less consensus here. There are staunch cognitivist theories on which emotions have cognitive content, much like or even exactly like beliefs. Martha Nussbaum, for example, argues that our emotions contain "judgments about important things" which involve "appraising

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 10
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

an external object as salient for our own well-being" (2001: 19). Non-cognitivist theories maintain that emotions lack cognitive content. Jesse Prinz, for example, holds that emotions are "somatic signals . . . not cognitive states" although they "represent concerns" (2007: 68). Moreover, while we often think of emotional processes as rapid and automatic, they can be more drawn out and consciously accessible. One can, for example, be acutely aware of one's anxiety and its bodily effects, which may ebb and flow over the course of days or weeks, as opposed to occurring in rapid episodes typical of fear or anger.

I suspect the concept of emotion is flexible and not amendable to precise definition. I'm certainly not fond of classical analyses of concepts, which posit necessary and sufficient conditions (May & Holton 2012; May 2014b). In any case, we can be ecumenical and conceive of emotions as mental states and processes that have certain characteristic features. Heidi Maibom provides a useful characterization of *emotions* as "mental states associated with feelings, bodily changes, action potentials, and evaluations of the environment" (2010: 1000; cf. also Haidt 2003: 853).

Suppose I negligently step on your gouty toe, so you become angry with me. Your anger has an *affective element*: a characteristic feel. The emotion also has *motivational elements* that often appear to activate relevant behavior: e.g., it motivates you to retaliate with verbal and physical abuse (but see Seligman et al. 2016: ch. 8). Emotions also seem to have *physiological effects*—e.g., your anger will lead to a rise in blood pressure, increased heart rate, and other bodily changes. Finally, feeling angry also typically involves or at least causes *cognitive elements*, such as thoughts about my blameworthiness, about the damage to your toe, about how you could best retaliate, and so on.

I will understand such *cognitive* elements as, roughly, mental items whose function is to accurately represent. A cognitive mental state, like a belief, can be contrasted with motivations, goals, or *desires*, which arguably function to bring about the state of affairs they represent (Smith 1994). Tim and I may both *believe* that there is a taco on the table, but only I *want* to eat it, for he is stuffed. My longing for the scrumptious taco involves a desire or a mental state whose function is to bring it about that I eat the taco. Importantly, cognitive elements represent how things are and can thus play a role in inference. Insofar as emotions can have cognitive elements or at least effects on cognition, emotions can provide information and facilitate reasoning.

The cognitive elements or effects of emotions make the apparent reason/emotion dichotomy blurry at best. Despite the similarities between the two, however, at least one important difference may remain: it's commonly assumed that *feelings* are essential to emotions but not to the process of

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 11

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

reasoning. Many researchers use the term "affect" to refer to a kind of feeling (see, e.g., Seligman et al. 2016: 50), although it is something of a technical term with different meanings for some theorists. Perhaps, then, we should just speak of the dichotomy between inference/affect or cognitive/non-cognitive states. However, sometimes the connection to rationalism and sentimentalism is clearer if we operate with the working conception of reasoning and emotion and then contrast their cognitive vs. affective aspects.

So far, the working conception respects the worry that there is no sharp division between reason and emotion. This *overlap view*, as we might call it, seems to satisfy many in empirical moral psychology (e.g., Greene 2008; Maibom 2010; Helion & Pizarro 2014; Huebner 2015). For others, however, it doesn't go far enough.

On the *total collapse view*, there is no difference between reasoning and emotional processing. Peter Railton, for example, construes the "affective system" quite broadly such that "affect appears to play a continuously active role in virtually all core psychological processes: perception, attention, association, memory, cognition, and motivation" (2014: 827; cf. also Damasio 1994; Seligman et al. 2016). On this picture, it may seem that the debate between rationalists and sentimentalists is spurious, since affect and inference are inextricable. However, what motivates the collapse view is a range of empirical evidence which suggests that "emotion" turns out to be more like inference than we thought, not that "reason" turns out to be less like inference than we thought. As James Woodward has put it, areas of the brain associated with emotion are "involved in calculation, computation, and learning" (2016: 97).

This would be a welcome result for the view to be defended in this book, which aims to emphasize the role of reasoning and inference in moral psychology. Indeed, the affective system broadly construed is something humans share with many other animals (Seligman et al. 2016). The total collapse view suggests that affective processes are necessary for moral judgment merely because they're *required for inference generally*, moral or otherwise. So we give sentimentalists a better chance if we operate with the overlap view instead. To see this, we need to consider in more detail the debate between rationalists and sentimentalists.

1.2.3 Rationalism vs. sentimentalism

Clearly, both reason and emotion play a role in moral judgment. Nevertheless, a traditional dispute remains between rationalists and sentimentalists over the comparative roles of *inference vs. feelings* in distinctively moral cognition (Nichols 2008: n. 2; Maibom 2010: 1000; May & Kumar forthcoming). The issue

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 12
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

is interesting in its own right and we'll eventually see that it has important practical implications for how to develop moral knowledge and virtue.

The empirical claim made by those in the *rationalist* tradition is that reasoning is central to moral cognition in a way that the affective elements of emotions are not. Such (empirical) rationalists hold that moral judgment, just like many other kinds of judgment, is fundamentally "a product of reason" (Nichols 2004: 70) or "derives from our rational capacities" (Kennett 2006: 70). However, as only a psychological thesis, "rational capacities" here is meant to be non-normative—even poor reasoning counts as deriving from one's "rational" capacities. We can more clearly capture this idea by construing rationalism as the thesis that moral judgment is ultimately "the culmination of a process of reasoning" (Maibom 2010: 999). Emotions can certainly influence moral cognition, according to rationalists, but primarily insofar as they facilitate inference; they aren't essential for making a judgment distinctively moral.

On the *sentimentalist* picture I'll resist, mere feeling or the affective component of emotions is essential for moral cognition and thus moral knowledge (if such knowledge is possible). Without emotions, a creature can't make any moral judgments, because the feelings constitutive of emotions are in some way essential to having moral concepts. As Hume famously put it, when we condemn an action or a person's character:

> The vice entirely escapes you, as long as you consider the object. You never can find it, till you turn your reflection into your own breast, and find a sentiment of disapprobation, which arises in you, towards this action. (1739–40/2000: 3.1.1)

Hume clearly conceives of such sentiments or passions as feelings, and it's this aspect of emotions, not their role in inference, that sentimentalists see as distinctive of moral judgment. Contemporary sentimentalists, such as Shaun Nichols, continue this line of thought, stating that "moral judgment is grounded in *affective* response" (2004: 83, emphasis added). Moreover, sentimentalists don't merely claim that lacking feelings or affect would hinder moral judgment, but rather that this would render one incapable of understanding right from wrong. Even when sentimentalists emphasize the importance of reasoning and reflection in moral judgment, they remain sentimentalists because they give "the emotions a constitutive role in evaluative judgment" in particular (D'Arms & Jacobson 2014: 254; cf. also Kauppinen 2013).

Rationalists can agree that emotions are commonly involved in human moral judgment and that lacking them leads to difficulties in navigating the social world. Humans are undoubtedly emotional creatures, and sentiments pervade social interactions with others. To build a moral agent, one might have to endow

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 13

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

it with emotions, but only because a finite creature living in a fast-paced social world requires a mechanism for facilitating rapid reasoning and quickly directing its attention to relevant information. A creature with unlimited time and resources needn't possess emotions in order to make distinctively moral judgments (cf. Jones 2006: 3).

On the rationalist view, the role of emotions in morality is like the role of ubiquitous technologies: they facilitate information processing and structure our way of life. If the Internet was somehow broken, for example, our normal way of life would be heavily disrupted, but it's not as though the Internet is fundamental to the very idea of communication and business transactions. Of course, in one sense the Internet is essential, as we rely on it for how we happen to operate. But a cognitive science of how communication fundamentally works needn't feature the ability to use email. No doubt the analogy only goes so far, since emotions are not some recent invention in human life. They are part of human nature, if there is such a thing. The point is simply that, for sentimentalists, emotions are more than vehicles for information processing; they partially define what morality is. Thus, even if emotions aid in reasoning, we still can conclude that their affective elements aren't necessary for moral judgment. The sentimentalist tradition isn't vindicated if emotions are merely ways of processing information more quickly, rigidly, and without attentional resources (see Prinz 2006: 31).

Of course, emotions may be required for moral judgment, especially knowledge, merely because experiencing certain emotions seems necessary for knowing what another is feeling. Indeed, sentimentalists sometimes draw an analogy between moral judgments and judgments about color: they are both beliefs typically caused by certain experiences (e.g., Hume 1739–40: 3.1.1; Prinz 2007: 16; Slote 2010; Kauppinen 2013: 370; Sinhababu 2017: ch. 4). The relevant experience may then be necessary for knowledge, particularly because such experiences are conscious, or essentially qualitative, mental states. And understanding what a sensation or experience is like seems impossible without having it oneself (Jackson 1982). In the moral domain, men in power have historically taken a paternalistic attitude toward women, and yet men presumably don't generally know exactly what it's like to be a woman or to carry a child to term. As some liberals are fond of saying: If men were giving birth, there wouldn't be much discussion about the right to have an abortion. Perhaps even women don't know these things either until they have the relevant experiences (see Paul 2014). Similarly, an emotionless robot may be ignorant of some moral facts in virtue of lacking feelings of love, grief, pride, or fury.

Even so, this doesn't show that emotions are essential for making a moral judgment. At best, certain experiences are sometimes required for understanding

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 14
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

a phenomenon. A sophisticated robot could acquire the relevant knowledge by having the requisite experiences. In fact, this is just an instance of a more general problem of ignorance of morally relevant information. Suppose I visit my grandmother in the hospital in Mexico. I know what it is to suffer but I falsely believe that the Spanish word "sufre" refers to, not suffering, but the vegetarian option at a Chipotle restaurant. Then I won't know that the nurse did wrong when she made "mi abuela sufre." Does this imply that Spanish is essential for moral knowledge? In certain circumstances, I must know the relevant language, but this is too specific for a general characterization of what's psychologically essential for moral judgment. Similarly, suppose one doesn't fully understand, say, the anguish of torture or the humiliation of discrimination unless one experiences them first-hand. Such examples don't demonstrate that feelings are essential for making distinctively *moral* judgments but rather judgments about specific cases. The theoretically interesting position for sentimentalists to take is the one that many have indeed taken: emotions are required for understanding right from wrong generally, not merely for understanding a subset of particular moral claims.

1.3 Pessimism about Moral Motivation

1.3.1 Sources of pessimism

Suppose the previous challenges have been rebutted: ordinary moral cognition is a fundamentally rational enterprise capable of rising to moral knowledge or at least justified belief. Still, we might worry that we rarely live up to our scruples, for self-interest and other problematic passions too frequently get in the way. Even if we do end up doing the right thing, we do it for the wrong reasons. When we're honest, fair, kind, and charitable, it's only to avoid punishment, to feel better about ourselves, or to curry someone's favor. Something seems morally lacking in such actions—let's say that they're not fully *virtuous*. Just as merely true but unjustified belief doesn't seem to deserve a certain honorific (e.g., "knowledge"), merely doing the right thing, but not for the right reasons, doesn't warrant another moniker ("virtue").

To be truly virtuous, it seems in particular that moral considerations should more frequently guide our behavior; reason cannot be a slave to non-rational passions, selfish or otherwise. Kant (1785/2002) famously thought that only such actions—those done "from duty"—have moral worth. For example, we'd expect a virtuous merchant not only to charge a naïve customer the normal price for milk but to do it for more than merely self-interested reasons—e.g., to avoid a bad reputation.

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 15

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

Many believe the science warrants pessimism: deep down we're primarily motivated to do what's right for the wrong reasons, not morally relevant considerations. Robert Wright, for example, proclaims that an evolutionary perspective on human psychology reveals that we're largely selfish, and yet we ironically despise such egoism:

> [T]he pretense of selflessness is about as much a part of human nature as is its frequent absence. We dress ourselves up in tony moral language, denying base motives and stressing our at least minimal consideration for the greater good; and we fiercely and self-righteously decry selfishness in others. (1994: 344)

This disconcerting account paints us as fundamentally egoistic. On the most extreme version—*psychological egoism*—all of one's actions are ultimately motivated by self-interest. We are simply incapable of helping others solely out of a concern for their welfare. An ulterior motive always lurks in the background, even if unconsciously.

There is a wealth of rigorous research that seems to suggest that altruism is possible particularly when we empathize with others. However, compassion can be rather biased, parochial, and myopic. We are more concerned for victims who are similar to ourselves, or part of our in-group, or vividly represented to tug at our heartstrings, rather than a mere abstract statistic (Cialdini et al. 1997; Jenni & Loewenstein 1997; Batson 2011). Moreover, studies of dishonesty suggest that most people will rationalize promoting their self-interest instead of moral principles (Ariely 2012). Even if we're not universally egoistic, we may not be far from it (Batson 2016).

A related source of pessimism draws on the vast research demonstrating the *situationist* thesis that unexpected features of one's circumstances have a powerful influence on behavior. Many have taken this literature to undermine the existence of robust character traits or conceptions of agency and responsibility that require accurate reflection. However, even if we jettison commitments to character traits and reflective agency, results in the situationist literature pose a further challenge. If our morally relevant actions are often significantly influenced by the mere smell of fresh cookies, the color of a person's skin, an image of watchful eyes, and the like, then we are motivated by ethically arbitrary factors (see, e.g., Nelkin 2005; Nahmias 2007; Vargas 2013; Doris 2015). A certain brand of situationism, then, may reveal that we're chronically incapable of acting for the right reasons.

Suppose we do often do what's right for more than self-interested or arbitrary reasons. Proponents of *Humeanism* would argue that, even when we behave morally, we are beholden to our unreasoned passions or desires (e.g., Sinhababu 2009;

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 16
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

Schroeder, Roskies, & Nichols 2010). If Humeans are right, our actions are always traceable to some ultimate or intrinsic motive that we have independent of any reasoning or beliefs. Bernard Williams famously discusses an example in which a callous man beats his wife and doesn't care at all about how badly this affects her (1989/1995: 39). On the Humean view, we can only motivate this man to stop his despicable behavior by getting him to believe that being more kind will promote something he already cares about. We must try to show him that he'll eventually be unhappy with himself or that his treasured marriage will fall apart. Pointing out that he's being *immoral* will only motivate if he happens to care, and care enough, about that. If, however, refraining from physical abuse will not promote anything this man already wants, then the Humean says there is nothing that could motivate him to stop except a change in his concerns.

The Humean theory can be conceived as a kind of pessimism if acting for the right reasons requires ultimately acting on the basis of recognizing the relevant reasons, not an antecedent desire. Some, like Thomas Reid, seem to think so:

> It appears evident . . . that those actions only can truly be called virtuous, and deserving of moral approbation, which the agent believed to be right, and to which he was influenced, more or less, by that belief. (1788/2010: 293)

We do often describe one another's actions this way—e.g., "She did it because she knew it was the right thing to do"—without appealing to an antecedent desire to be moral.

However, Humeans might retort that acting for the right reasons requires only being motivated by specific moral considerations (e.g., kindness, fairness, loyalty), not the bare belief that something is right per se (cf. e.g., Arpaly 2003: ch. 3). Perhaps, for example, a father shouldn't have "one thought too many" about whether he should save his own drowning daughter over a stranger's (Williams 1976/1981). In general, the virtuous person presumably wouldn't "fetishize" morality but rather be ultimately concerned with the welfare of others, fidelity to one's commitments, and so on (Smith 1994), and a moral belief might still be problematic in this way (Markovits 2010). We'll grapple with this issue later (Chapters 7–8), but for now suffice it to say that a certain kind of pessimism about the role of reason in moral motivation remains if Humeanism is right.

For a variety of reasons, pessimists conclude that the aim of doing what's right for the right reasons is practically unattainable. On a common account of what's required for virtuous motivation, it's out of reach for most of us. I aim to show that we are capable of genuinely altruistic motivation and that our beliefs about what we ought to do can motivate action without merely serving or furthering some antecedent desire. Moreover, while features of the situation

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 17

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

certainly influence what we do, the ethically suspect influences do not systematically conflict with virtuous motivation. I ultimately argue that humans are capable of acting from duty or doing the right thing for the right reasons. Morally good motives are not rarities.

1.3.2 Non-cognitivism & relativism

The discussion so far has assumed that we can have moral beliefs, conceived as distinct from emotions, desires, or other passions. A complete defense of anti-Humeanism and rationalism requires showing that moral judgments don't just express non-cognitive states. Consider, for example, the sentence "Slavery is immoral." It seems such sentences don't always merely express one's negative feelings toward slavery. That is, it seems that *non-cognitivism* about moral judgment is false. Unlike beliefs, mere feelings and desires arguably can't be evaluated for truth or accuracy, which makes it difficult to see how they can be part of a process of reasoning or inference.

Importantly, rejecting non-cognitivism needn't commit one to denying *relativism*, the view that moral statements are only true relative to some framework, such as the norms of one's culture. I don't assume that moral judgments are robustly objective but rather that they can be cognitive, similar to other beliefs. When I say, "Lebron is tall," this may be true only relative to a certain contrast class (ordinary people, not basketball players), but it is nonetheless assessable for truth or falsity in a certain context. In a somewhat similar fashion, moral truths are nonetheless truths even if they are in some sense relative to a culture, species, or kind of creature. So we needn't assume that moral truths are objectively true—a core element of *moral realism* (Shafer-Landau 2003)—in order to defend moral knowledge, conceived as justified true belief.

I don't intend to argue at length against non-cognitivism. The view has largely already fallen out of favor among many researchers. A survey of philosophers conducted in 2009 reveals that only 17 percent lean toward or accept it (Bourget & Chalmers 2014: 476). There is good reason for this. The famous Frege-Geach problem, which I won't rehearse here, shows that non-cognitivists struggle to make sense of moral language without drastically revising our best conception of logic and semantics (Schroeder 2010). Non-cognitivism is not exactly a live empirical theory either, as psychologists and neuroscientists appear to assume that moral judgments express beliefs. For example, rather than simply identify moral judgments with emotions or desires, researchers look to whether emotions are a cause or consequence of the moral judgment. In fact, the vast majority of "pessimists" I'll target assume cognitivism as well.

Moreover, we needn't accept non-cognitivism to account for the various uses to which moral judgment can be put. For example, *hybrid theories* can capture the idea that we sometimes use sentences like "That's just wrong" to express a negative reaction, like a feeling or desire, or to express a belief that an action or policy is wrong. Compare statements containing a pejorative, such as "Yolanda's a Yankee," which in some countries is used to express both a belief (Yolanda is American) and a distaste for her and other Americans (Copp 2001: 16). I favor something like this model (May 2014a), according to which moral judgments can express both cognitive and non-cognitive states (cf. also Kumar 2016a). However, I assume here only the falsity of non-cognitivism, which is compatible with either a hybrid view or a strong cognitivist theory on which moral judgments only or chiefly express beliefs.

1.4 Optimistic Rationalism

My primary aim is to resist the predominant pessimism about ordinary moral psychology that has developed in light of scientific research on the human mind. I will offer a more optimistic defense of ordinary moral thought and action in which reason plays a fundamental role—*optimistic rationalism*, if you will.

Since pessimism comes in many forms, an optimistic view must be multi-faceted, with various components in opposition to the variety of pessimistic arguments. In particular, I aim to undermine some popular sources of empirically grounded pessimism (see Figure 1.1). I thus contend that moral judgments are generated by fundamentally cognitive and rational processes (rationalism), which are not subject to wide-ranging empirical debunking arguments (anti-skepticism). Moreover, moral motivation is not always ultimately egoistic (psychological altruism), is heavily driven by a concern to do what's right, and is not always a

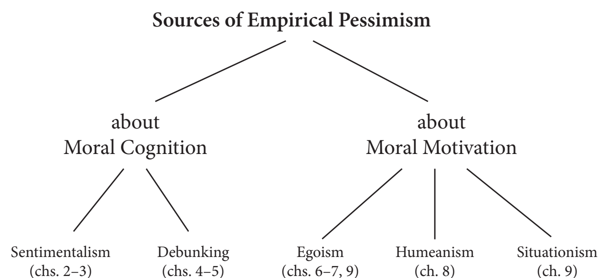

Figure 1.1. Key Sources of Empirically Grounded Pessimism
Note: Parentheses indicate in which chapters the source is primarily addressed.

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 19

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

slave to unreasoned passions (anti-Humeanism). All of this casts doubt on the idea that virtuous motivation is rare among ordinary individuals (anti-skepticism).

Some may regard this cluster of views as closely associated with the Kantian tradition in moral philosophy. However, one can defend an optimistic picture of moral psychology without adopting a specific Kantian articulation of what precisely makes an action immoral. For example, Kant (1785/2002) says an action is wrong if the maxim on which it is based can't be rationally chosen as a universal law. The theory developed in this book does not commit to knowledge of such specific accounts of fundamental moral principles. It's similar in some important respects to the moral psychology of the great Chinese philosopher Mencius (Morrow 2009) and of some contemporary philosophers who are not particularly Kantian. So non-Kantian moral theorists—especially virtue ethicists, but even some consequentialists—may find much to agree with in what follows.

At any rate, few optimists have taken the empirical challenges seriously, let alone answered them successfully. Some valiant attempts are simply incomplete in that they only address one aspect of moral psychology, such as moral judgment (e.g., Maibom 2005; Kamm 2009; Kennett & Fine 2009; Mikhail 2011; Sauer 2017) or moral motivation (e.g., Kennett 2002; Kennett & Fine 2008; de Kenessey & Darwall 2014; Sie 2015). Others claim to be optimists but embrace what I regard as sources of pessimism, such as simple sentimentalism (e.g., de Waal 2009) or revisionary utilitarianism (e.g., Greene 2013). This book aims to provide a more complete and satisfactory defense.

I employ a divide and conquer strategy, breaking our moral minds into two key components (and their corresponding normative ideals): moral judgment (and knowledge) and moral motivation (and virtue). Consider how these two may come together or apart. Suppose you're deciding whether you ought to press charges against your thieving son who is in the grips of a severe drug addiction. If all goes well, you form the correct judgment, it's warranted or justified, and you thus know what to do. Suppose you decide it's best to proceed with the charges. Next is the important task of living up to your standards. If you're virtuous, you will act according to this judgment and for the right reasons, yielding moral motivation that exhibits virtue.

One of my overarching aims is to reveal the deep connections and parallels in these two aspects of our moral minds—judgment and motivation—which are often addressed separately and by different sets of researchers. In subsequent chapters, we'll see that our moral beliefs are formed primarily by on the basis of unconscious inference, not feelings, and that these moral beliefs play a prominent role in motivating action.

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 20
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

1.4.1 From moral judgment to knowledge

The next four chapters form Part I, which tackles moral judgment and to what extent it rises to knowledge or at least justified belief.

Chapter 2 ("The Limits of Emotion") argues that, contrary to the current sentimentalist orthodoxy, there is insufficient reason to believe that feelings play an integral role in moral judgment. The empirical evidence for sentimentalism is diverse, but it is rather weak and has generally been overblown.

Chapter 3 ("Reasoning beyond Consequences") turns to some of the complex inferential processes that do drive ordinary moral thinking. Ample experimental evidence establishes in particular that we often treat more than just the consequences of one's actions as morally significant. Ultimately, much of ordinary moral judgment involves both conscious and unconscious reasoning about outcomes and an actor's role in bringing them about.

But don't we have empirical reasons to believe that core elements of ordinary moral judgment are defective? Chapter 4 ("Defending Moral Judgment") argues that ordinary moral cognition can yield justified belief, despite being partly influenced by emotions, extraneous factors, automatic heuristics, and evolutionary pressures. I rebut several prominent, wide-ranging debunking arguments by showing that such pessimists face a Debunker's Dilemma: they can identify an influence on moral belief that is either defective or substantial, but not both. Thus, wide-ranging empirical debunkers face a trade-off: identifying a substantial influence on moral belief implicates a process that is not genuinely defective.

By restoring reason as an essential element in moral cognition, the foregoing chapters undermine key sources of support for the sentimentalists and the debunkers. Such pessimists have tended to accept the idea that feelings play an important role in ordinary moral judgment. Sentimentalists embrace this as a more or less complete characterization. Debunkers instead use the apparent power of emotion as a source of skepticism about either all of moral judgment or only some of its more intuitive bases. With a regard for reason, ordinary moral thinking is on safer ground.

However, while moral knowledge is possible, Chapter 5 ("The Difficulty of Moral Knowledge") admits that we are far from flawless moral experts. There are two key empirical threats to the acquisition or maintenance of well-founded moral beliefs. First, empirical research can indeed reveal questionable influences on our moral views. While wide-ranging debunking arguments are problematic, this does not hinder highly targeted attacks on specific sets of moral beliefs (e.g., some influenced by implicit biases). Second, while people share many values, most ordinary folks have foundational disagreements with others who are just as likely

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 21
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

to be in error ("epistemic peers"). However, this threat is likewise constrained since many moral disagreements aren't foundational or aren't with what most people should regard as their peers.

1.4.2 From moral motivation to virtue

Part II consists of four chapters that focus on ordinary moral action and whether it's compatible with virtuous motivation, which involves doing the right thing for the right reasons.

Chapter 6 ("Beyond Self-Interest") argues that we can ultimately be motivated by more than egoistic desires. Decades of experiments in social psychology provide powerful evidence that we are capable of genuine altruism, especially when empathizing with others. The psychological evidence, moreover, cannot be dismissed as showing that empathy blurs the distinction between self and other so much that it makes helping behavior non-altruistic.

Even if we can rise above self-interest, we may just be slaves to our largely, if not entirely, egoistic passions. Chapter 7 ("The Motivational Power of Moral Beliefs") argues that the motivational power of reason, via moral beliefs, has been understated. A wide range of experimental research shows that when we succumb it's often due in part to a change in moral (or normative) belief. Rationalization, perhaps paradoxically, reveals a deep regard for reason—to act in ways we can justify to others and to ourselves. The result is that, even when behaving badly, actions that often seem motivated by self-interest are actually ultimately driven by a concern to do what's right (moral integrity). This addresses a second form of egoistic pessimism but also sets up a challenge to the Humean theory addressed in the next chapter.

Chapter 8 ("Freeing Reason from Desire") picks up on the idea that our beliefs about which actions we ought to perform have a pervasive effect on what we do. Humean theories would of course insist on connecting such beliefs with an antecedent motive, such as a desire to do what's right. However, I first shift the burden of proof onto Humeans to motivate their more restrictive, revisionary account. I then show that Humeans are unlikely to discharge this burden on empirical grounds, whether by appealing to research on neurological disorders, the psychology of desire, or the scientific virtue of parsimony.

Chapter 9 ("Defending Virtuous Motivation") considers further empirical threats to our ability to act for the right reasons. There are two main threats: self-interested rationalization and arbitrary situational factors. However, wide-ranging versions of such empirical challenges resemble sweeping attempts to debunk moral knowledge, and they're likewise subject to a dilemma. One can easily identify an influence on a large class of actions that is either substantial or

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:50 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 22
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

defective but not both. Thus, like moral knowledge, the science suggests that the empirical threat to virtue is limited.

1.4.3 Moral enhancement

The previous chapters defend the idea that, based on our regard for reason, ordinary moral thought and action are capable of rising to knowledge and virtue. But of course, such optimism must be cautious. We do often behave badly, or do what's right for the wrong reasons, or lack justified moral beliefs.

Chapter 10 ("Cautious Optimism") serves as a brief conclusion with a recapitulation of the main claims and moves made in the book, along with a discussion of how moral knowledge and virtue can be enhanced. One broad implication of optimistic rationalism is that the best method for making more of us more virtuous will not target our passions to the exclusion of our cognitive, reasoning, and learning abilities. However, sound arguments aren't enough, for human beings are fallible creatures with limited attention spans. Still, the impediments to virtue are not primarily the absence of reason or our basic modes of moral thought; rather we must combat ignorance, self-interested rationalization, and the acquisition of misinformation and vices.

There is further reason for caution and caveat. For all I will say here, one might adopt a truly global skepticism and conclude, on empirical grounds, that we don't know right from wrong and can't act virtuously because reason itself is thoroughly saturated with defective processes, both inside and outside the moral domain. It's beyond the scope of this book to grapple with such a deep skepticism about our cognitive and inferential capacities. A vindication of moral knowledge or virtue, especially given a rationalist moral psychology, would ultimately require defending the reliability of our cognitive faculties generally. I'll be content here, however, if I can show that empirical research doesn't reveal that reason is largely absent or defective in our basic modes of moral thought and motivation.

1.5 Coda: Appealing to Science

We'll encounter a great deal of empirical research throughout this book. We should proceed with some caution given heightened awareness of concerns arising in experimental psychology and other sciences.

First, there is a somewhat surprising amount of *fraud*, in which researchers fabricate data—and moral psychologists are no exception (Estes 2012). Second, there is an unsettling amount of poor scientific practice. Much of this falls under the heading of *p-hacking*, as when researchers continuously run participants in

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:51 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 23

OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

a study until they find a statistically significant result, which increases the likelihood of a false positive. Third, the scientific process itself has flaws. For example, there are *publication biases* in favor of shocking results and against null findings, including failures to replicate a previous result. One consequence is the *file drawer problem* in which failures to detect a significant effect are not published or otherwise circulated, preventing them from being factored into the cumulative evaluation of evidence. Related to this, the rate of replication seems unfortunately low in the sciences generally, including psychology in particular—an issue some call *RepliGate* (e.g., Doris 2015). A recent group of over 200 researchers attempted to carefully replicate 100 psychological studies and found roughly that only 39 percent succeeded (Open Science Collaboration 2015).

An additional problem is that much of the empirical research in moral psychology is done on a small portion of the population, typically undergraduates in North American and European universities. That is changing, as researchers are increasingly recruiting participants from outside of universities, including some from multiple cultures. Still, as Joseph Henrich and his collaborators (2010) have put it, the majority of research participants are from societies that are predominantly Western, educated, industrialized, rich, and democratic (WEIRD people). This is especially problematic when we have empirical evidence that what appear to be psychological universals are not, at least not to the same degree in all societies.

Of course, we shouldn't overreact. The vast majority of scientists are not frauds and many conduct careful and rigorous studies. While participants are often WEIRD, such a subject pool may suffice if one's aim is merely to establish the existence or possibility of certain psychological mechanisms, not their universality. Moreover, replication attempts shouldn't necessarily be privileged over the original studies. The original could have detected a real effect while the later study is a false negative. The cutoff for statistical significance (typically $p < 0.05$) is somewhat arbitrary, after all. A statistically significant result only means, roughly, that there is a low probability (less than 0.05) that the observed difference, or a greater one, would appear in the sample, even when there is no real difference in the population (that is, when the null hypothesis is true). The p-value importantly doesn't represent the probability that any hypothesis is true but rather a conditional probability: the probability of observing a certain result assuming that the null hypothesis is true. Thus, if a replication attempt is close to passing the conventional threshold—nearly yielding a successful replication—we may still have some reason to believe in the effect. Observing a difference between experimental groups that yields a p-value of 0.06, for example, doesn't exactly amount to conclusive reason to accept the null. In general, it's more difficult to

Comp. by: SatchitananthaSivam Stage : Revises1 ChapterID: 0003404371 Date:30/1/18
Time:22:13:51 Filepath:D:/BgPr/OUP_CAP/IN/Process1/0003404371.3d
Dictionary : OUP_UKdictionary 24
OUP UNCORRECTED PROOF – REVISES, 30/1/2018, SPi

prove a negative (e.g., that an effect is bogus) than it is to establish the existence of a phenomenon.

There is certainly room for improvement in science, including larger sample sizes, more replication attempts, and more cross-cultural research. But science can clearly advance our knowledge, even about the mind and our complex social world, provided we aren't overly credulous. For example, as Machery and Doris (2017) emphasize, one shouldn't stake a conclusion on a single study, ideally not even on a few studies from one lab, especially when sample sizes are low. It's best to draw on a large set of studies in the literature, appealing where possible to meta-analyses and reviews, while recognizing of course that these aren't definitive either. Caution and care can ultimately yield strong arguments based on scientific data.

Despite judicious appeal to the science, I tread lightly when drawing conclusions from empirical studies or philosophical analysis. Like Hume, I suspect the truth about such perennial issues will be difficult to uncover, and "to hope we shall arrive at it without pains, while the greatest geniuses have failed with the utmost pains, must certainly be esteemed sufficiently vain and presumptuous" (173–940: intro, 3). So I don't claim to have conclusively proven the theses in this book. Thankfully, though, my main aim is more modest. Defending a more optimistic conception of our righteous minds requires merely showing that it's a plausible approach given our best evidence to date. No chapter is meant to establish definitively the relevant claim it defends. The value of the book is largely meant to arise from all of the parts coming together to exhibit a counterweight to the pessimistic trend.