

## NATURALIZING MORAL NATURALISM

Jessica Isserow

ONE OF the most pressing tasks for metaethicists is that of solving the *location problem*: finding a home for morality in the natural world. It goes without saying that some have risen to the occasion more enthusiastically than others, and it is one enthusiast in particular that shall occupy my attention here. The naturalist moral realist affirms continuity between ethics and the empirical sciences, striving to integrate her metaethics with the outputs of scientific theorizing. To her mind, moral epistemology does well to take science as its guide; moral facts are ripe for empirical investigation.<sup>1</sup>

Unfortunately, the naturalist canon does not always reflect these noble ambitions.<sup>2</sup> The naturalist is committed to letting the world do (much of) the talking. But so far, she has scarcely given it the chance to speak. My aim here is to set us back on course. The organizing theme of this paper is that the outputs of empirical investigations are of underrecognized significance for the moral naturalist. Its more specific contention is that these empirical resources help her to address two fundamental challenges that she faces.

Moral naturalists are often said to have trouble accommodating the *intensional* and *extensional* character of morality.<sup>3</sup> A metaethical position accommodates morality's *intensional* character just in case it is in keeping with (what are commonly regarded as) important conceptual commitments of moral thought and talk. Moral naturalism seems to fail dismally in this regard, for it is famously unfaithful to what many take to be a core conceptual commitment of moral discourse: that all agents have reason to act as morality requires independently of their contingent ends. Indeed, naturalists usually take an agent's reasons

- 1 For representative declarations of these commitments, see Boyd, "How to Be a Moral Realist"; Railton, "Naturalism and Prescriptivity." Different naturalists will admittedly embrace these commitments to different degrees (see section 2 below).
- 2 Hereafter, I substitute "(moral) naturalist" and "(moral) realist" for the more cumbersome "naturalist moral realist." There are obviously other sorts of naturalists and other sorts of realists, but they are not my focus here.
- 3 I borrow the distinction from Southwood (*Contractualism and the Foundations of Morality*), who uses it to assess different varieties of moral contractualism.

to be moral to be *hostage to* such ends. Naturalists' critics allege that this outlook fails to take morality seriously as a normative phenomenon. Call this the *intensional challenge*.<sup>4</sup>

A metaethical position accommodates morality's *extensional* character just in case it (largely) accords with substantive judgments concerning the extension of terms such as "morally required" and "morally impermissible"—for example, the judgment that it is morally impermissible to subject people to inhumane treatment on account of their skin color. Here again, the naturalist seems to come up short. This is because (as I explain below) her method for identifying which natural properties are (or constitute) the moral ones is fairly *permissive*; it seems to allow for moral truths that conflict with our substantive moral judgments. Of course, no metaethical theory can plausibly be expected to take *none* of these judgments to be mistaken or confused. Properly understood, then, the concern is not simply that the naturalist allows for moral truths that conflict with these judgments, but that she allows for moral truths that conflict with them in rather *striking* ways. Call this the *extensional challenge*.

As I conceive of these challenges, their upshot is as follows: the naturalist has incurred significant explanatory debt to date, and it is imperative that she either pay off this debt or discharge it. The naturalist could pay off her debt by demonstrating that she can indeed accommodate the intensional and extensional dimensions of morality. Alternatively, she could discharge her debt by establishing that the phenomena she fails to accommodate are not properly viewed as central to either dimension. To my mind, the naturalist has not exercised her full potential in either regard, for she is yet to fully avail herself of the resources at her disposal—insights from evolutionary theory, psychology, and ethnography in particular. I will argue that these resources help her to address both challenges in a more satisfying way. This is not to peddle the radical thesis that metaethics is a battle best fought on empirical ground. But it does, I think, demonstrate what we stand to gain by covering multiple terrains in our philosophical pursuits.

My first order of business will be to spell out the commitments of moral naturalism (section 1). I will then turn my attention to the intensional challenge (section 2). Here, naturalists have traditionally responded with an optimistic prognosis: given widespread human interests and concerns, many of us do as a matter of fact have reason to be moral—enough of us to vindicate intuitions concerning morality's normative credentials. No doubt, this prognosis has high

4 My locution (*the intensional challenge*) is admittedly misleading; this is certainly not the *only* sort of intensional challenge that naturalists face. Naturalists have also been pressed for not securing the right kind of objectivity, for example. See Loeb, "Gastronomic Realism"; Kurth, "What Do Our Critical Practices Say about the Nature of Morality?" I am picking my battles here.

intuitive plausibility. But it is only a start. Claims about human interests are, after all, *empirical claims*—and the naturalist’s claim still remains in need of sustained empirical attention. My own objective will be to defend this claim as a robust empirical generalization. In section 3, I tackle the extensional challenge. Many worry that the naturalist’s method for identifying which natural properties are moral properties commits her to classifying as morally permissible a range of behaviors that we regard as morally perverse. I shall argue that the naturalist can distinguish perverse moral frameworks from legitimate ones on principled grounds. Nothing I say entails that moral naturalism is home and dry. But my arguments do suggest that the position has far more going for it than many have thought.

### 1. MORAL REALISM, NATURALLY

A moral naturalist takes moral properties to be natural properties.<sup>5</sup> My discussion will be restricted to *realist* varieties of moral naturalism, according to which moral judgments are beliefs, there are moral properties, and these properties are constituted by or identical to mind-independent natural properties. The notion of mind independence can be tricky to tie down. For my part, I take it to be best captured by the thought that moral truths hold independently of our attitudes in the sense that they are not constituted by our beliefs or opinions about them.<sup>6</sup> It is trickier still to tie down the notion of a natural property. I will work with an understanding in which natural properties are those susceptible to empirical investigation.<sup>7</sup> Nothing of great import will hang on this understanding. I provide it in the interest of situating the discussion.

There are two distinguishing commitments of moral naturalism that will be relevant to my purposes. I do not claim that each is necessary to qualify as a member of the camp. In the absence of both, however, I think it is safe to say that the position in question would be one that parted ways from the majority of the naturalist canon. The first commitment is rooted in a particular approach to normativity. Traditionally, the naturalist takes an agent’s (intrinsic) desires to be the ultimate source of her reasons for action.<sup>8</sup> What fundamentally grounds

5 I intend this characterization to include nonreductive naturalists, who take moral properties to be natural properties that supervene on but are not reducible to other (nonmoral) natural properties.

6 See Brink, *Moral Realism and the Foundations of Ethics*, 18–20.

7 Railton, “Naturalism and Prescriptivity,” 154; Copp, “Why Naturalism?” 185. Cf. Sturgeon, “Moore on Ethical Naturalism,” 538.

8 I am here skirting around different (and no doubt more sophisticated) articulations of this commitment. Rather than speak of an agent’s (intrinsic) desires, the naturalist may prefer

an agent's normative reasons—and, in turn, her (ir)rationality—is what she fundamentally cares about. When paired with the naturalist's characterization of moral facts themselves, this approach yields a contingent relationship between moral requirements and an agent's normative reasons. Moral facts are mind-independent natural facts that bear no essential connection to an agent's desires. Whether or not morality provides her with normative reasons is thus completely hostage to whether she cares about the ends at which morality aims, or the moral good as such.<sup>9</sup>

A word of caution here. To claim that morality is not necessarily a source of normative reasons is not to claim that morality is not necessarily a source of *moral* reasons. The naturalist can readily admit that there are moral reasons in the sense that there is a system of requirements that sanctions morally good behavior. But this is not saying much. (There are also reasons of etiquette in this sense.) In denying that morality is necessarily a source of normative reasons, the naturalist is denying that moral requirements have intrinsic reason-giving force—denying that any agent would be irrational, or at least guilty of a normative mistake, were she unresponsive to them.<sup>10</sup> Obviously, none of this entails that the naturalist will look upon moral failures approvingly. She will, however, resist describing them as *rational* failures. Henceforth, I shall refer to this commitment as *no necessary irrationality in immorality*.

Second, the naturalist is committed to a certain permissiveness concerning the matter of which natural properties are the moral ones—though some care must be taken in spelling out the precise sense of permissiveness at issue. To this end, it is helpful to consider the naturalist's recipe for identifying which natural properties are moral properties. Often, she begins with the observation that morality has one or more distinctive functions in human life—making flourishing societies possible or stabilizing cooperation, for example.<sup>11</sup> She then draws upon these in sorting the moral from the nonmoral. Moral codes are said to be those that best serve a “society's needs and non-moral values.”<sup>12</sup>

---

to say that it is her needs and values or her ends that ground her normative reasons. (See Copp, *Morality, Normativity, and Society*, ch. 9; Railton, “Moral Realism.”) Either could be substituted for “intrinsic desires” without obscuring the basic point.

- 9 Railton, “Moral Realism,” 166; Brink, *Moral Realism and the Foundation of Ethics*, ch. 3; Boyd, “How to Be a Moral Realist,” 340–42.
- 10 These remarks reflect the well-known distinction between normativity in the rule-involving and reasons-providing senses. See Parfit, *On What Matters*, 2:267–68.
- 11 See Copp, *Morality, Normativity, and Society*; Sterelny and Fraser, “Evolution and Moral Realism.”
- 12 Copp, *Morality, Normativity, and Society*, 159–60.

Moral facts are said to be facts about “human cooperation and the social practices that support [it].”<sup>13</sup>

Claims such as these reflect claims about which natural properties the moral properties are (or are likely to be) given the roles they are usually taken to play. Sometimes, such claims are presented as the product of conceptual analysis.<sup>14</sup> Other times, they are put forward in the spirit of an empirical hypothesis. This is a choice point that marks a divide between *a priori* and *a posteriori* varieties of moral naturalism. For my purposes, I need not take sides. What it is important to appreciate is that on neither outlook do these claims decide all first-order moral issues in advance. It remains an open question what the moral truths actually are—which moral code does as a matter of fact best promote human flourishing or cooperation, say. What the moral truths actually are remains an open empirical question, one that ultimately hinges upon what the world turns out to be like—let the empirical chips fall where they may. I will refer to this latter commitment as *the open-endedness of morality*.<sup>15</sup>

Although both *no necessary irrationality in immorality* and *the open-endedness of morality* are widely embraced among moral naturalists, they lead to two

13 Sterelny and Fraser, “Evolution and Moral Realism,” 984. Cf. Boyd, “How to Be a Moral Realist,” 329.

14 Jackson, *From Metaphysics to Ethics*.

15 Some may question my suggestion that *a priori* naturalists take morality to be open ended in this way. Jackson takes a conceptually competent, idealized reasoner with full information about the world to be in a position to know the moral truths (“From Metaphysics to Ethics,” 31–42). (“Full information” must be qualified lest the reasoner’s knowledge be trivial. We might restrict it to information about the worldly supervenience base described in some semantically neutral language.) Given this, it may be difficult to see how *a priori* naturalists such as him are faithfully characterized as viewing moral truths as open ended. Surely our moral concepts, together with the state of the world, *already* decide what the moral truths actually are? But concerns such as these merely require that we precisify the sense of open-endedness at issue. For Jackson, the moral truths depend upon our network of implicit or explicit moral beliefs and opinions (“folk morality”) following critical reflection (“mature folk morality”) *together with* the state of the world (“From Metaphysics to Ethics,” 133). The conceptually competent idealized reasoner can determine which natural facts the moral facts *would be* at each world that is presented to her for consideration. But she cannot know which natural facts the moral facts *actually are* until she knows precisely which world she is in—until, that is, she knows (at least some, perhaps all) empirical facts as well. Moral truths are therefore open ended even for the *a priori* naturalist, in the sense that their content crucially hangs upon the nature of a particular class of nonmoral facts. Both *a priori* and *a posteriori* naturalists are thus committed to the kind of open-endedness that derives from leaving the fate of moral truths in the hands of some yet-to-be-determined, nonmorally specified state of the world. Crucially, as we will see in section 3, it is this sense of the open-endedness of morality that naturalists’ opponents capitalize upon.

well-known challenges. The naturalist's commitment to *no necessary irrationality in immorality* invites the accusation that she fails to do justice to morality's normative credentials. Given her commitment to *the open-endedness of morality*, she has also been taken to task for overgenerating admissible moralities. These are, respectively, the intensional and extensional challenges for moral naturalism. I take each in turn.

## 2. THE INTENSIONAL CHALLENGE

The naturalist who occupies my attention is committed to *no necessary irrationality in immorality*. Her critics allege that this outlook fails to take morality seriously as a normative phenomenon. Taking morality seriously (so this line goes) requires taking it to be intrinsically reason giving. The naturalist must, however, view this as an exercise in vaulting ambition. She simply cannot take morality seriously if this is what taking morality seriously requires.

My aim in what follows will be to show that the intensional challenge loses much of its sting once we turn our attention to deep-seated features of human sociality and psychology. I begin by distinguishing different faces of the problem (section 2.1). Ultimately, I think the naturalist should concede to her opponents that there is a deep connection between moral requirements and reasons for action. She should, however, deny that it is deep in the way they think it is. More specifically, the naturalist should take "human beings have reason to be moral" to be true when construed as a *robust empirical generalization*.<sup>16</sup> Though not all agents take morality seriously, *we* certainly do—and it turns out to be surprisingly difficult for us not to. The task for section 2.2 will be to defend the empirical plausibility of this generalization. Defending it as a solution to the intensional challenge will be the business of section 2.3.

A caveat will be helpful before proceeding. Some philosophers may find themselves puzzled by *normativity*—not just by *moral normativity*. Some variation of the intensional (and indeed, extensional) challenge that I address here likely arises for naturalism about the normative as well as moral naturalism. And while certain lessons may carry over, I do not pretend to be offering a comprehensive response to the broader suite of challenges that have been brought to bear against naturalist approaches in metaethics. The paper is addressed, then, most directly to those philosophers who find themselves especially puzzled by normativity in the moral sphere—though I do hope

16 As I have noted, many naturalists accept (something like) this empirical claim. Yet they seldom if ever build a convincing empirical case for it.

those whose puzzlement extends wider still will be able to extract something useful from it as well.

### 2.1. *Interpreting the Challenge*

The intensional challenge to moral naturalism has many faces. In the interests of focusing the discussion, I want to isolate two strands of thought within this family and concentrate my critical attention there. The first takes the naturalist to task for identifying morality in something *external* to human agents.<sup>17</sup> Insofar as the naturalist regards moral properties as mind independent—something “out there” waiting to be discovered—her proposal gives rise to the potential for a gaping distance between the moral facts and the ends with which we identify. It remains an open question whether anyone *cares* about these facts. Yet it does not seem to be an open question whether we care about morality. Moral considerations have a deep practical hold upon us. The naturalist, then, appears to be looking for morality in all the wrong places. If we are to close the normative gap that she leaves wide open, then we would likely do better by understanding morality in terms of something *internal* to human agency—perhaps even in terms of human agency itself. This face of the intensional challenge reflects a concern about *normative distance*.

A second face draws upon deep-seated intuitions concerning morality’s normative reach. Many find it highly intuitive that *all* agents have reason to be moral—not merely those who have particular desires or ends that would be served by being moral. This intuition seems to favor the proposal that moral reasons are *categorical* in character: that moral requirements provide an agent with normative reasons *independent* of whatever desires or preferences she happens to have.<sup>18</sup> More often than not, the point is driven home by calling upon a variety of morally dubious characters ranging from opportunists to outsiders. One well-known opportunist is Hume’s sensible knave, who only acts morally well when doing so is to his benefit.<sup>19</sup> Outsiders are different: their psychological architecture differs in rather fundamental ways from our own. They and their ilk lie at the outskirts of the human community. (Extreme sadists are an exemplary case.) What opportunists and outsiders have in common is a desire set that is best served by acting contrary to moral requirements. Given this, the naturalist seems committed to saying that they lack reason to be moral. Yet that assessment fails to respect the intuition that these individuals *do* have reason to be moral. Opportunists and outsiders seem guilty of a normative mistake;

17 Korsgaard, *The Sources of Normativity*, 112.

18 Shafer-Landau, “A Defence of Categorical Reasons.”

19 Hume, *Enquiry Concerning the Principles of Morals*.

they do not merely appear to ignore rules that it is optional to take seriously. Morality's normative reach thus seems more extensive than the naturalist can allow inasmuch as she constructs our reasons to be moral upon contingent foundations such as human preferences. This second face of the intensional challenge reflects a concern about *normative jurisdiction*.

In addressing the intensional challenge, I propose to draw inspiration from Hume. When reflecting upon morally dubious characters, Hume concedes that such persons may lack the sorts of concerns that yield reasons to be moral. Speaking of his sensible knave, he remarks: "if his heart rebel not against such pernicious maxims, if he feel no reluctance to the thoughts of villainy or baseness, he has indeed lost a considerable motive to virtue."<sup>20</sup> But to this, Hume adds an important qualification. Though the knave's heart may not rebel against pernicious principles, *ours* certainly do. Most people *are* emotionally constituted and socially situated such that they have reason to be moral.

Hume's strategy takes us some way in addressing concerns about morality's normative jurisdiction; if *most* human agents have reasons to be moral, then morality's jurisdiction is fairly respectable.<sup>21</sup> Though Hume's reasoning may seem simple and straightforward, I think it bears more fruit than one might expect. Indeed, a little empirical digging reveals it to be capable of mitigating concerns about normative distance as well. I do not doubt that there are a number of respectable interpretations of Hume's argumentative strategy. But I am going to propose that we develop his insights along the following lines: "human beings have reason to be moral" is true when construed as a *robust empirical generalization*. Let me explain what I mean by this using an analogy.

Consider the following generalization *E*: "elephants care for their young." This is true when interpreted as a statistical claim, for *most* elephants do care for their young. But it can also aspire to be more than a statistical claim, for there is a deeper explanation for this statistic; it is something about *the nature of elephants* that *explains why* they care for their young. Elephants mature slowly, relying on maternal (and allomaternal) milk for nutrition during their early years. Mothers also usually give birth to one calf at a time, a conservative reproductive strategy that favors high investment in individual offspring. Thus, *E* is also true when interpreted as a *characterizing generic*, a claim that—as I am understanding it here—tells us something about what is *normal* for members of a kind.<sup>22</sup> For our purposes, we can understand normality in terms of deeply

20 Hume, *Enquiry Concerning the Principles of Morals*, sec. 9, pt. 2.

21 It remains to be shown whether it is *respectable enough*. I take up this concern in section 3.3.

22 See Nickel, "Generics and the Ways of Normality." Cf. Cavedon and Glasbey, "Outline of an Information-Flow Model of Generics."

entrenched features of a kind's members. These are features on which many others depend, and which, if changed, would result in a wholesale change in the kind. Such features need not be strictly speaking intrinsic; elephants' high degree of sociality, for instance, is plausibly counted among them.

To be sure, *E* admits of exceptions. Elephant mothers in circuses and zoos, for example, reject or kill their calves surprisingly often—something that is partly accounted for by many of these mothers having been deprived of close contact with older females themselves when they were young.<sup>23</sup> But such exceptions are no threat to the truth of *E* as an empirical generalization. Insofar as they are exceptions, they seem to be principled ones: they are exceptions that prove the rule. If we wanted to transpose this idea into a framework of characterizing generics, then we might build upon Nickel's suggestion that "normality" always depends to some degree upon our "inductive target" as well as features of the kind in question.<sup>24</sup> If the rearing habits of elephants are our inductive target, then our interest presumably concerns what their typical caring practices look like. Yet caring practices—like much else in biology—reflect "a complex interaction" between a range of factors.<sup>25</sup> Some of these factors will be less relevant than others given our theoretical aim of arriving at useful and informative empirical generalizations about elephant rearing. Facts about elephants' maturation cycle and reproductive strategy, for instance, seem relevant given this aim, whereas influences introduced by human circuses and zoos do not.

My goal in what follows will be to argue that "human beings have reason to be moral" (henceforth, I will call this *H*) is on a par with *E* in these respects. *H* is plausibly true when interpreted as a statistical claim, a claim about what holds true for *most* of us. But it is *also* true when interpreted as a characterizing generic, a claim that reflects what is normal for us and holds true in virtue of deeply entrenched features of our psychology and sociality. (As we shall see, this is an instance where the evidence that favors the generic claim favors the statistical claim as well.) As I hope these qualifications make clear, the ensuing discussion is simply intended to illuminate phylogenetically ancient and important features of human beings that ground their reasons to be moral—no suspicious normative or teleological assumptions are being smuggled in. My arguments neither presuppose nor are intended to support any notions of natural human goodness and defect or standards of human excellence.<sup>26</sup>

23 Kurt and Mar, "Neonate Mortality in Captive Asian Elephants (*Elephas maximus*)."

24 Nickel, "Generics and the Ways of Normality," 643–45.

25 Nickel, "Generics and the Ways of Normality," 643.

26 Cf. Foot, *Natural Goodness*; Hurka, *Perfectionism*.

To my mind, much of the foregoing has to some extent been lost in previous responses to the intensional challenge. To be sure, it has not been lost on naturalists that humans typically have reason to be moral.<sup>27</sup> Yet it has not been sufficiently emphasized that this claim can aspire to be more than a mere statistic. Naturalists do, to their credit, sometimes appeal to facts about the human condition—to our sympathy or sociality, for example.<sup>28</sup> But that is only a start. It is no substitute for engagement with the rich empirical literature bearing on the matter. This literature reveals that our reasons to be moral have deep psychological roots.

Let me add one final clarification prior to proceeding. One may wonder whether, given *H*, humans generally have *decisive* reason to be moral (such that being moral is always rationally required) or merely *sufficient* reason to be moral (such that being moral is always rationally permissible). Speaking of “decisive” reasons here may seem a little ambitious. I would be content if my arguments at least supported the conclusion that humans often have decisive reason and very often sufficient reason to be moral. (Indeed, some may well find this less ambitious conclusion closer to the truth!) I would not be content if my arguments merely suggested that humans generally have some (perhaps vanishingly small) reason to be moral. Though I wager that the latter claim is true, it is not enough to address the intensional challenge.

Those skeptical of my strategy will fall into two camps. Some will suspect that *H* is false. Others will insist that, even if *H* is true, it does little to mitigate the intensional challenge. I will tackle each skeptic in turn.

## 2.2. *Is the Empirical Generalization Plausible?*

The naturalist who occupies my attention is, recall, committed to a particular conception of normative reasons: what grounds an agent’s normative reasons is (roughly) what she cares about. Insofar as the naturalist takes normative reasons to be rooted in an agent’s conative psychology, then, the task of supporting *H* is effectively the task of showing that moral edicts have strong resonance with us in a manner that reflects entrenched features of human psychology. I will focus upon three features in particular: our prosocial emotions, our sociality, and our need for good repute. As will become clear, there is a great deal of explanatory overlap here, for these features are mutually supporting.

Beginning with the prosocial emotions, humans clearly care about the welfare of others. Even human infants exhibit strong other-regarding concerns.<sup>29</sup>

27 See Railton, “Moral Realism,” 170; Copp, *Morality, Normativity and Society*, 244.

28 Boyd, “How to Be a Moral Realist,” 341; Copp, *Morality, Normativity and Society*.

29 Liszkowski et al., “12- and 18-Month-Olds Point to Provide Information for Others”; Warneken and Tomasello, “Altruistic Helping in Human Infants and Young Chimpanzees.”

This is unsurprising. The survival of our species has long been predicated upon successful cooperation, and there has long been biological and cultural selection for emotional responses that support it. We feel sympathy in response to others' suffering, anger in response to their transgressions, and guilt in response to transgressions of our own.

Each of these experiences has attentional and motivational import. Prosocial emotions focus our attention on others—on their needs, their actions, and their situations. Empathetic emotions ground other-regarding concerns and motivate helping behavior.<sup>30</sup> Anger is a formidable motivator too: as its intensity increases, so too do the costs we are willing to incur to penalize mistreatment.<sup>31</sup> Guilt plays a central role in maintaining interpersonal relationships, urging us to repair social bonds that are threatened by our misdemeanors.<sup>32</sup> It should be emphasized that these experiences are not merely influential—they are typically quite powerful. Even *proactive* guilt can rein in a temptation to renege on social commitments.<sup>33</sup>

Some may complain that the behavior these emotions (dis)incentivize is merely *prosocial*—that it is not yet *moral*. But prosocial emotions need not fly solo. In human social worlds they are governed by shared standards and expectations. Norms *direct* our feeling; they tell us where to focus our sympathy, how much anger is warranted, and whether and when guilt is appropriate. There is a fundamental social need to coordinate our behavior, and a tried and trusted way of doing so is to direct our prosocial emotions toward similar action classes; to feel anger and guilt (in appropriate measure) in response to the same transgressions, and to reserve empathy for the same sorts of people. Moral norms, then, *direct* our prosocial responses in ways that build upon our emotional architecture as well as our capacities for rational reflection—not just any instance of guilt or empathy is sanctioned. Moral education is in part education in how to feel.<sup>34</sup>

With that said, it is worth observing that a healthy correlation between prosocial and moral (that is, morally sanctioned) behavior is precisely what many varieties of naturalism predict. Recall that for (many) naturalists, what

30 Eisenberg et al., "Relation of Sympathy and Personal Distress to Prosocial Behavior"; Findlay et al., "Links between Empathy, Social Behavior, and Social Understanding in Early Childhood"; Toi and Batson, "More Evidence That Empathy Is a Source of Altruistic Motivation."

31 Bosman and van Winden, "Emotional Hazard in a Power-to-Take Experiment"; de Quervain et al., "The Neural Basis of Altruistic Punishment."

32 Tangney and Dearing, *Shame and Guilt*, 124–25, 184–85.

33 See Frank, *Passions within Reason*.

34 Mameli, "Meat Made Us Moral."

distinguishes morality from other normative domains is precisely the *kinds* of natural phenomena that it makes its business: “society’s needs,” “human cooperation,” or facts about “harm” and “benefit.”<sup>35</sup> Again, this is not to suggest that being prosocial is *coextensive* with being moral. It is only to point out that a healthy degree of overlap here is to be expected. Reflection, deliberation, and negotiation have significant roles to play in the formation of moral norms as well.<sup>36</sup>

Importantly, the relationship between norms and emotion cuts both ways. Moral norms not only direct human feeling but have been molded by it. Every culture’s moral package builds upon our affective architecture.<sup>37</sup> Norms backed by feeling gain more traction; emotion makes particular standards more salient and memorable.<sup>38</sup> Our emotional configuration also constrains the norms that we can get behind. Norms that align more closely with our affective predispositions tend to be more learnable. None of this is to deny cross-cultural variation. Different packages of norms build upon different features of human psychology, and they do so in different ways.

The prosocial emotions therefore enable coordination and promote cooperative response. Both features are important. The satisfaction of most of our human needs (such as subsistence and security) depends in some way upon our social group. Members of human societies have long been interdependent, and as a result their survival has long depended upon effective cooperation.<sup>39</sup> Our psychology reflects this history. We are adapted for interactive and collaborative living, cognitively as well as emotionally. Humans value joint activities intrinsically. Human children not only value cooperative games that lack an instrumental rationale but often transform tasks with an instrumental aim into cooperative interactions.<sup>40</sup>

35 See, respectively, Copp, *Morality, Normativity, and Society*, 159–60; Sterelny and Fraser, “Evolution and Moral Realism,” 984; Foot, “Moral Arguments,” 510.

36 See Kitcher, *The Ethical Project*. A related worry about my arguments is that they only establish reasons to conform to social conventions, not reasons to be moral. In fact, however, they establish both. The considerations discussed here explain why people often have strong incentives to conform to group norms. In addition, however—and as the considerations introduced in section 4 will help to drive home—they also explain why people often have strong incentives to defy them when they are judged or felt to be wrong, or when they conflict with powerful prosocial impulses that favor kindness or mercy. Any view that denied that both incentives are typically present would have a hard time accounting for social change born of moral resistance—and why such change often meets resistance in turn.

37 Haidt, *The Righteous Mind*.

38 Nichols, *Sentimental Rules*.

39 Tomasello, *A Natural History of Human Morality*.

40 Tomasello, *Why We Cooperate*, 63–65.

In a social world predicated upon cooperation, it helps to have someone to cooperate *with*. It is especially helpful to have a *good* cooperative partner—one who will not leave you for dead as soon as the cooperative labor has borne its fruit. This insight features prominently in evolutionary accounts of morality, where theories of partner choice often play a central role.<sup>41</sup> But the basic lesson carries over to the present. Nice guys do not necessarily finish last; *ceteris paribus*, they tend to do rather well on the cooperative market. We want to mix in the same circles as the reliable, the trustworthy. Even human infants prefer those with helpful dispositions.<sup>42</sup>

Reputation matters in a cooperative market. Having a bad name means having less social capital—no one wants to be paired up with a knave. And being short on collaborators carries real costs. The price of being unpopular is high, ranging from lower job prospects to lower life expectancy.<sup>43</sup> There are internal costs as well. We feel shame when others think ill of us. And shame is highly punitive.<sup>44</sup> It is bad—and it *feels bad*—to get a bad rap. Indeed, social disapproval is often regarded as an especially toxic form of punishment. Many report preferring pain, incarceration, amputation, or even *death* to a heavily tarnished reputation.<sup>45</sup> Humans care about how they fare in the court of public opinion. We take active steps to shape our reputations, and not only by acting in socially sanctioned ways but also “by joining in the conversation” about our actions and justifying them to others.<sup>46</sup>

The foregoing strongly suggests to me that human beings are generally emotionally and socially situated such that they have reason to be moral. Strong other-regarding concerns, a high degree of interdependence, and a need for others’ good opinion are widespread and deeply entrenched features of our psychology and ways of life. Given this, characteristically moral behavior (helping others, say) tends to be to our benefit. Importantly, these insights would seem to apply to Hume’s knave as well. Our social preferences favor authentically helpful dispositions—not opportunism. And social acceptance is not a luxury the average person can afford to forgo. The satisfaction of many fundamental

41 Tomasello, *A Natural History of Human Morality*; Stanford, “The Difference between Ice Cream and Nazis.”

42 Kuhlmeier, Wynn, and Bloom, “Attribution of Dispositional States by 12-Month-Olds”; Hamlin, Wynn, and Bloom, “Social Evaluation by Preverbal Infants.”

43 Western, Kling, and Weiman, “The Labor Market Consequences of Incarceration”; House, Landis, and Umberson, “Social Relationships and Health.”

44 Tangney and Dearing, *Shame and Guilt*, 137–38.

45 Vonasch et al., “Death before Dishonor.”

46 Sperber and Baumard, “Moral Reputation,” 511.

human needs—alliances, romantic partnerships, careers—depends upon the accumulation of social capital.

Of course, the knave has a response at the ready here: no one need *know* that he is only virtuous for show. So long as he *appears* virtuous, he can avoid the costs that accompany a bad reputation. Yet this seems to reflect a naïve optimism on his part. The proposal that we are capable of systematically fooling others about our moral caliber lacks empirical plausibility. People are rather good at predicting others' cooperative intentions, especially those of acquaintances.<sup>47</sup> All in all, pretending to be good is a dangerous game.

Simply put, the best way to earn a good reputation is to deserve it—to actually *be* good.<sup>48</sup> At a minimum, being good requires internalizing standards: developing dispositions to feel anger when such standards are violated, and guilt when one falls short. We have seen that these prosocial emotions have motivational value. But they confer *signaling* value as well; emotional response is hard to fake, making it a fairly reliable sign of moral commitment.<sup>49</sup> These assurances are important, for we do not only choose collaborators with an eye to their track record—we try to make reasonable inferences about their mental states as well.<sup>50</sup> It is for this reason that the policy of behaving morally only when morality pays is not an especially promising policy. The kind of moral behavior that tends to pay is the kind that stems from sincere commitment.<sup>51</sup>

So much for the knave. What of the outsider, though? The outsider, recall, lacks typical human motives. She is (let us suppose) indifferent to others' opinions, impervious to guilt, and perfectly content to go it alone. Some will be inclined to regard outsiders as unassailable counterexamples to *H*. On reflection, however, I think it is more plausible to view them as principled exceptions

47 See Brosig, "Identifying Cooperative Behavior"; Frank, Gilovich, and Regan, "The Evolution of One-Shot Cooperation"; Pradel, Euler, and Fetchenhauer, "Spotting Altruistic Dictator Game Players and Mingling with Them."

48 Frank, *Passions within Reason*; Sterelny, *The Evolved Apprentice*, ch. 5; Sperber and Baumard, "Moral Reputation."

49 Frank, *Passions within Reason*.

50 See Sperber and Baumard, "Moral Reputation," 507. I should add that this is nothing approaching the whole story. Moral emotions likely differ in their signaling value; some may be less difficult to fake than others. See O'Connor, "The Evolution of Guilt." And I do not mean to claim that we make judgments about others *purely* on the basis of their emotional profile. More plausibly, we collate and draw upon different sources of evidence—reputation, behavior, emotional response—in arriving at a judgment.

51 At this stage, some readers will want to object that these observations only demonstrate (at best) that we have reasons to *be* good. To this, they will be quick to add that the reasons to be moral that the intensional challenge demands are reasons to *do* good. Rest assured, I take up this challenge in section 3.3.

to it. As should now be clear, a callous disregard for others is hardly characteristic of human beings. Indeed, this is among the key diagnostic criteria for a range of human *pathologies*.<sup>52</sup> Thus, psychopaths and other populations with systematic deficiencies in affective response pose no threat to our empirical generalization. These are the easy cases, the exceptions that prove the rule.

Or are they? Perhaps the easy cases are not *quite* so easy. One complicating factor is that disorders such as psychopathy lie within a spectrum, and that not all members of this population lie at its extreme end. Those members who do will invite the response above. But what of those who merely show psychopathic tendencies? The further away someone lies from the extreme, the *less* likely it becomes that the person will *lack* characteristic other-regarding concerns or the need for social support networks. But then, it also becomes *more* likely that the person will have reasons to be moral—in which case, we can simply apply the same reasoning we applied to the knave. The “principled exception” response, then, primarily concerns extreme outsiders; I do not deny that this phenomenon is graded in important ways.<sup>53</sup>

Let me now consider what I take to be the hard cases. There have long been drastic inequalities in wealth and power within human societies. Those who enjoy disproportionate shares of these resources (whom I will simply refer to as “elites”) are not uncharacteristic human beings in the manner that outsiders are; they presumably share the same characteristic human concerns as the rest of us. But this would seem to spell trouble for *H*, for it is not clear that elites *do* have reasons to be moral. The features of our psychology and social lives that I have emphasized—other-regarding concerns, reputation, interdependence—seem far less pronounced in those occupying the upper echelons of society. Elites surely need not invest so much in their reputation; they seem less entrenched in our networks of interdependence and less beholden to others than the rest of us.

For my part, I think that elites *do* have reasons to be moral. It is true that Elon Musk does not have to cozy up to his boss to get a promotion. However, he is not practically isolated from the rest of society either. Stock prices in Musk’s companies have been known to plummet as a result of his careless words (he once called an analyst a “boring bonehead”). Of course, such interdependence is plausibly more common in modern liberal societies. (Louis XIV certainly did not have to

52 American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*.

53 Yet another complication is that not all naturalists may agree with my suggestion that psychopaths in particular qualify as outsiders. Some, of course, will: my treatment above is importantly similar to Boyd’s, for instance (though unlike him, I do not trace the psychological abnormality back to a “cognitive deficit.”) See Boyd, “How to Be a Moral Realist,” 340–42. But other naturalists seem to favor a contrary perspective. See, for instance, Brink, “Responsibility, Incompetence, and Psychopathy.”

worry about his stock prices falling.) Bullying may therefore be a safer strategy in certain kinds of environments. Still, it is not without risk; radically discounting others' interests carries dangers of its own.<sup>54</sup> (I elaborate upon these in section 3.)

Though the above perspective strikes me as reasonable, let me offer a further possibility for those unconvinced. Elites may very well qualify as highly *uncharacteristic* human beings in a perfectly good sense. Here, we might recall Nickel's suggestion that "normality" is determined not merely by features of a kind but by our inductive target as well—something that, in turn, is informed by our interest in arriving at useful empirical generalizations.<sup>55</sup> If it is the factors underlying humans beings' reasons to be moral that are our inductive target, then our theoretical interest presumably concerns what human beings' typical ways of relating to one another look like. To this end, reputational effects, emotional dispositions, and social interdependence are relevant factors to consider. But the peculiar position of elites—who are socially situated in strikingly different ways to just about everyone else given their comparative lack of interdependence—arguably are not.<sup>56</sup>

My treatment of the hard cases thus has two prongs. I myself believe there is a reasonable case to be made that elites *do* have reason to be moral. That said, there is also an alternative outlook according to which such persons simply fall outside the scope of our inductive target and lack reason to be moral. Either way, there is no trouble for *H*.

### 2.3. *Is It Enough?*

I have argued that human beings have reason to be moral, where that is understood as a robust empirical generalization. The task of the previous section was to motivate the empirical plausibility of that generalization. The task of this section will be to motivate its metaethical serviceability. I now address both faces of the intensional challenge.

Let me begin with the normative distance worry, which takes the naturalist to task for locating morality in something *external* to human agents. Insofar as

54 Elites such as Musk arguably have further interests that are served by being moral. A referee helpfully observes that this seems especially true when we turn our attention to moral norms of a broadly Kantian kind and focus on personal relationships. If Musk does not keep his promises or treat people with respect, for instance, then he presumably will not have many (true) friends.

55 Nickel, "Generics and the Ways of Normality," 642.

56 Of course, I have maintained throughout that *H* is true *both* as a characterizing generic *and* as a statistical claim about most human beings. But the latter interpretation is easily accommodated in the case of elites. It is immensely difficult to reduce practical dependence upon others without certain resources (including vast sums of money and human capital). Clearly, not everyone has such resources, for not everyone can be in the top 1 percent.

moral facts are mind independent, there is thought to be a real potential for a gaping distance between them and the ends with which we identify. Upon reflection, however, this is not quite right. In light of the considerations raised in section 2.2, it is more perspicacious to view the naturalist as providing an answer that is at once internal *and* external. Moral facts are not simply facts about our wills or attitudes; they are objective facts about the practices that effectively support our cooperative endeavors. In this sense, they are external. But *which* practices effectively support our cooperative endeavors depends heavily upon deep-rooted features of human psychology. As we have seen, not all norms are created equal; those that have greater motivational uptake are more likely to be preserved and passed on. The prosocial emotions thus establish a harmony between motivation and moral response. Moral facts, then, have a crucial *internal* element, for they depend in crucial ways upon our emotional constitutions.

I turn now to the normative jurisdiction worry: On my account, is morality's normative jurisdiction extensive enough? I have conceded that outsiders lack reason to be moral, and have argued that this need not undermine *H*. Still, some may worry about outsiders escaping our *criticism*. Insofar as we concede that outsiders lack reason to be moral, we seem to have rendered inadmissible any normative complaint we might have had against them. We can no longer charge such individuals with having failed to acknowledge the reasons within their normative landscape. If they care not for moral matters, and we acknowledge that there is no reason for them to do so, then what is left for us to say to them?

I am inclined to view this question as premised upon a faulty assumption—namely, that we must have something *to say to* the outsider. We will certainly want something *to do about* them. And we will certainly have a lot *to say about* them. Yet it is difficult to see why they are properly viewed as a target of moral conversation. One can understand why opportunists meet this condition. (Here, the call is coming from inside the house!) But the outsider is incapable of authentic participation in moral life. Perhaps she is someone to be controlled or contained—but she is surely not someone to be *convinced*. For a genuine conversation to proceed, there must be common ground—something we clearly lack with the outsider.<sup>57</sup>

It is worth saying something more in defense of this position, especially since some may take the treatment of outsiders to crosscut the intensional and extensional challenges. (If we deny that outsiders have reason to be moral, then we may risk not getting the content of morality quite right either.) In response, it bears mentioning, following Sharon Street, that the naturalist can

57 Cf. Manne, "Internalism about Reasons," 96–97, 103; Woods, "Footing the Cost (of Normative Subjectivism)."

easily accommodate our *morally unfavorable opinion* of the outsider.<sup>58</sup> We are indeed saying something true when we describe the outsider as a cad, an evildoer, a villain, or a malefactor (or various other things that should probably be omitted from an academic philosophy paper). Of course, such accusations are unlikely to strike fear into the heart of the outsider. But they certainly hit a nerve *with us*. Given the enormous importance that we attach to our terms of interaction with one another, being branded a moral pariah is just about the deepest insult we have at our disposal.

The opponent of naturalism's worry, then, cannot be that the naturalist is incapable of charging the outsider with an important normative failing. The worry is that the naturalist cannot charge the outsider with *a particular kind* of normative failing: namely, a failure to recognize and respond to reasons that she does indeed have. But the foregoing considerations should, I think, leave us feeling less confident that the latter charge is truly needed. Failing to recognize or respond to your own reasons is indeed a kind of shortcoming, one that signals some sort of normative defect, such as irrationality. But it is often regarded an even *greater* shortcoming to fail to live up to moral standards. Being a fool might be bad, but being a jerk is arguably worse.

Still, it would be helpful if the naturalist were capable of explaining away the intuition that outsiders *are* guilty of the normative failing that her opponent has in mind: Why is there the temptation to think that outsiders *do* have reason to be moral—when in fact they do not? My own suspicion is that this is likely owing to a common but understandable error in our thinking about them. Ordinarily, when we want to interpret others' behavior, we proceed on the assumption that they are fundamentally *like us*.<sup>59</sup> Outsiders, however, are *not* fundamentally like us. It is, then, rather misguided to import our own psychology into our efforts to interpret them—indeed, outsiders may well be unintelligible from our perspective.<sup>60</sup> But it is also understandable that we make such an error. Outsiders do not, after all, tend to announce their presence. And insofar as *H* holds true as a statistical claim as well, the default assumption that those outsiders we encounter are likely to have reasons to be moral seems justified, even if ultimately false.<sup>61</sup>

It follows from what I have said that morality's normative jurisdiction is not *limitless*—some human agents lie beyond its reach. But it does not necessarily

58 Street, "In Defense of Future Tuesday Indifference," 293.

59 See Davidson, "Radical Interpretation."

60 Notice that unlike outsiders, opportunists do seem intelligible, for they *are* fundamentally like us. And it is precisely *because* they are sufficiently like us that they typically *do* have reason to be moral.

61 Cf. Street, "In Defense of Future Tuesday Indifference," 293–94.

follow that I have failed to address the intensional challenge. Given the above arguments, it strikes me that we should now be far less confident that the folk view of morality finely distinguishes—let alone unambiguously decides—between the possibilities that reasons to be moral are:

1. Somewhat common among human agents, given their contingent preferences
2. Widespread among human agents, given their contingent preferences
3. Widespread among human agents in contingent albeit robust ways
4. Possessed by all human agents
5. Necessarily possessed by all human agents

Given my arguments, the naturalist can establish that morality's jurisdiction is respectable (possibility 3), even if it is not quite as expansive as some of us believed it to be (as supporters of possibilities 4 or 5 would have it).

But have I really *addressed* the opponents' concerns, or have I simply *dismissed* them? That opponent, recall, is not only concerned with capturing *a sizable contingent* of agents in our moral-reasons net; she is also concerned with *the manner* in which we do so. Following a well-known naturalist tradition, we have responded to the intensional challenge by appealing to widespread but *contingent* facts about human preferences. But what truly lies at the heart of the intensional challenge, it seems, is the concern that naturalism fails to capture the way in which moral reasons are *special*. And according to those who raise the challenge, this specialness is best spelled out in terms of such reasons being inescapable or categorical—that is to say, in terms of their *not* being premised upon contingent human preferences. Recall that it is precisely in virtue of making this move that the naturalist's opponent seems to do better justice to intuitions concerning morality's normative reach; they are able to say that its jurisdiction reflects something more in the order of possibility 5 in the list above.

In response, I think that the naturalist can readily concede that moral reasons *are* special. What she will deny is that they are special in the particular way her opponents take them to be. For the naturalist, what makes moral reasons special is in essence the special role they occupy in our hearts; as an empirical matter, human beings attach immense importance to being good and doing good. It is for precisely this reason, moreover, that we understandably (though, if I am right, mistakenly) take outsiders to have reasons to be moral: in our efforts to interpret and relate to them, we assume that they are fundamentally like us in this way.

Nevertheless, the opponent will understandably press: our reasons to be moral are still left hostage to ordinary human preferences on this approach. Are they not then rather *un-special*? Well, yes—if we have already decided in advance that being special requires being categorical. But one point that

needs emphasizing in this context is that we are not forced to choose between possibility 1 and possibility 5; something in between may well be capable of explaining the special importance we attach to the moral dimensions of our lives *and* vindicating our expectation that many (even if not all) others will tend to attach special weight to these dimensions as well. The naturalist, then, need not rest content with the claim that our reasons to be moral stand or fall with some *flimsy* alliance of human preferences. She can and should go further than this, emphasizing that many of the social and emotional factors that ground our reasons to be moral reflect robust features of our social environments and psychology. To put the point in a slogan: it need not follow from the fact that our reasons to be moral are contingent that they are *precarious*; some things are both contingent *and* robust.<sup>62</sup>

Let me conclude my discussion of the intensional challenge by considering a further feature of my arguments that some may take issue with. It may be objected that I have not shown that it is rational *to act* as morality recommends. At best, I have shown that it is rational to be (or become) *a person* who is intrinsically motivated to act as morality recommends. But surely it is the first of these conclusions that is needed to address the intensional challenge—not the second.

There is of course a respectable strategy for responding to this concern: forge a connection between the two conclusions. David Gauthier is well known for proposing that insofar as it is rational to be a virtuous person, it is also rational to perform acts that are the output of a virtuous disposition—even when doing so is not to one's immediate benefit.<sup>63</sup> While I think this response is on the right track, I do not wish to borrow from it too uncritically. Gauthier's project is driven by an ambition to account for morality's normative reach in a way that makes little if any appeal to other-regarding concerns. I do not share this ambition. Gauthier's strategy also comes dangerously close to simply redefining "rational action." Far from establishing a nexus between doing good and being good, one may worry whether the strategy is not too quick to simply assume or stipulate that there is one. Let me, then, do a little more to motivate the idea that there is such a nexus. I will proceed on the assumption that my arguments have established reasons to *be* good. The present question is how that could establish reasons to *do* good as well.

Notice first that becoming a good person usually involves reordering one's priorities. As a morally mediocre individual, I might value my career above

62 Though she does not quite put the point in these terms, I take it that something like this is what Foot was getting at in her discussion of the volunteers in the siege of Leningrad. See Foot, "Morality as a System of Hypothetical Imperatives," 310–11.

63 Gauthier, *Morals by Agreement*, 170–77.

all else, with my wardrobe coming in at a close second, and my family a distant third. After having undergone a process of moral character development, however, my preference ordering would likely be different. (A good person presumably puts their family before their wardrobe.) Part and parcel of being a morally good person, then, is having particular priorities—priorities that plausibly favor acting as morality requires on any (or many a) choice occasion. Although the morally mediocre may have preference orderings that justify neglecting their family for their careers, morally good people usually will not.

Further, it is not implausible that doing good may be important for *remaining* good. Moral action may be habit forming. Good behavior cultivates good character, as moral motives are reinforced by positive feedback from the social environment. Conversely, acting immorally often involves setting aside human feeling (pangs of guilt, say), as well as characteristic human concerns (“What will others think?”). The more we set aside such concerns, the more adept we are likely to become at overcoming the prosocial impulses that promote moral response. Overcoming these impulses is difficult, but it is by no means impossible. Thus, patterns of bad behavior likewise seem habit forming; over time, we risk breeding insensitivity. Doing bad may *make* us bad. At the very least, it seems apt to make us *less good*.

### 3. THE EXTENSIONAL CHALLENGE

Recall that a metaethical position accommodates morality’s *extensional* character just in case it (largely) accords with substantive judgments regarding the extension of moral terms such as “morally impermissible.” To see why the naturalist has trouble meeting this requirement, we can begin by revisiting her recipe for discovering which natural properties are the moral ones: she identifies some *nonmoral* purpose that morality serves and proposes that the moral facts are those that fit the bill. Consider, for example, Kim Sterelny and Ben Fraser’s contention that

a natural notion of moral truth falls out of the picture that moral belief evolved (in part) to recognize, respond to, promote, and expand the practices that make stable cooperation possible. For there are objective facts about the conditions and patterns of interaction that make cooperation profitable, and about those that erode those profits.<sup>64</sup>

On this approach, we are empirically corrigible when it comes to what the moral facts are—the naturalist lets the world do much of the talking. There is,

64 Sterelny and Fraser, “Evolution and Moral Realism,” 985.

however, a problem with letting the world do the talking: we might not like what it has to say. The naturalist proposes to single out moral properties by the nonmoral purposes they serve. Yet when the empirical dust has settled it may well turn out that these purposes are served in unsettling ways. There may be moral systems that fulfil morality's function (to promote cooperation, say) but sanction a range of behaviors that seem obviously morally impermissible. Insofar as the naturalist is committed to viewing these systems as comprising moral *truths*, she cannot plausibly accommodate morality's extensional character. It is this challenge that will occupy my attention in the remainder of this paper.

Two clarifications will be useful before proceeding. First, I am going to restrict myself in what follows to the variety of moral naturalism defended by Sterelny and Fraser.<sup>65</sup> This restriction is purely for illustrative purposes; the basic strategy could be enlisted by other naturalists as well. With that said, Sterelny and Fraser's framework seems especially likely to raise the following concern: Am I in the business of defending *realism* or *relativism* here? What if moral system<sub>1</sub> turns out to best promote cooperation for society<sub>1</sub>, whereas moral system<sub>2</sub> best promotes cooperation for society<sub>2</sub>? Many naturalists are, as it turns out, prepared for this eventuality. Frank Jackson argues that there is reason to expect convergence on a particular human morality but admits that this cannot be known in advance; if divergence truly is in our stars, then we should be willing to retreat into relativism.<sup>66</sup> Likewise, Richard Boyd thinks it is "pessimistic" to expect more than one human morality to emerge but concedes the possibility. (He nevertheless maintains that were it to eventuate, that would only "refute moral realism as that doctrine is ordinarily construed" and "would not undermine a generally realistic conception of moral language.")<sup>67</sup> But is this really all the relativism-realism divide boils down to: a mere empirical conjecture or a potentially misplaced hope? Perhaps so, perhaps not. I certainly do not want to pick *that* meta-metaethical battle here. The point is simply this: if the illustrative example raises concerns about realism retreating into relativism, it is arguably not unrepresentative of naturalism in this respect.

Moving on to our second clarification, it seems optimistic to expect that any specific naturalist identification of moral properties with natural properties is correct *as it currently stands*. Given this, we should not expect any present variety of naturalism to escape the extensional challenge *completely* unscathed. My ambition, then, is not to show that one promising implementation of naturalism is *completely immune* to the extensional challenge. It is rather to demonstrate that

65 Sterelny and Fraser, "Evolution and Moral Realism."

66 Jackson, *From Metaphysics to Ethics*, 137.

67 Boyd, "How to Be a Moral Realist," 351–52.

it is far more *resistant* to the challenge than is commonly thought. This will, I hope, make us less inclined to think the naturalist project is doomed to fail and more confident in its future. Although we are yet to develop a foolproof variety of moral naturalism, we are perhaps not quite so far off as some may fear.

Now to the demonstration. Sterelny and Fraser propose to understand moral truths as “maxims that are members of near-optimal normative packages—sets of norms that if adopted, would help generate high levels of appropriately distributed, and hence stable, cooperation profits.”<sup>68</sup> Their proposal is premised upon a particular empirically well-motivated picture of the evolution of human cooperation. To summarize, Sterelny and Fraser maintain that cooperative arrangements are more likely to be stable when the distribution of cooperative profits is *fair*—roughly, when there is not a huge disparity between any individual’s investment and her returns. When everyone is guaranteed roughly proportionate returns, everyone has a stake in the venture being successful. On this account, our moral psychology evolved to support effective cooperative arrangements such as these. We are adapted to “recognize, respond to, promote, and expand the practices that make stable cooperation possible.”<sup>69</sup>

A notable worry with this picture is that there seems to be nothing to prevent normative packages of the kind that interest Sterelny and Fraser from being morally *perverse*. Sterelny and Fraser may be forced to embrace the uncomfortable conclusion that what turn out to be the moral truths—for them, maxims that are members of a near-optimal normative package—conflict in striking ways with a swathe of substantive moral judgments. The worry is not baseless, especially given the details of Sterelny and Fraser’s proposal. Moral norms have a bad track record. The catalog of morally prescribed behaviors in human societies is dreadful, ranging from honor killings to foot-binding, female genital mutilation, and slavery.

At this juncture, a naturalist seems to find herself in a double bind. She cannot simply *define* the problem away. Building substantive moral premises into the conditions for effective cooperation amounts to abandoning her purely descriptive recipe for identifying the moral among the natural. Nor, it seems, can she hope to dismiss aberrant moralities on purely empirical grounds. If maxims that permit enslavement serve a society’s cooperative purposes, then the naturalist seems committed to viewing them as moral truths. Yet that seems wrong. As Max Barkhausen observes, “Most of us are deeply opposed to the idea that any way of coordinating on mutually beneficial behavior that our

68 Sterelny and Fraser, “Evolution and Moral Realism,” 985.

69 Sterelny and Fraser, “Evolution and Moral Realism,” 985.

moral evolution might have led us to endorse is as good as any other.”<sup>70</sup> As it turns out, however, the naturalist is *not* committed to the idea that any way of coordinating is as good as any other. On closer inspection, perverse moral norms are not particularly effective in promoting stable, efficient cooperation.

The thought that the naturalist is forced to accept whatever evolution throws her way seems to be premised upon an unreasonable optimality assumption. The assumption seems to be that a norm’s very existence as a facilitator of cooperation entails that it is part of a near-optimal normative package. This assumption is empirically suspect; in practice, many factors make moral optima difficult to reach. In what follows, I draw attention to some of the mechanisms that lead to the establishment and entrenchment of suboptimal normative packages. I will then explain why these packages are properly viewed as suboptimal—why they plausibly fail to promote stable cooperation. The mechanisms that I explore overlap to some degree. But each raises considerations distinct enough to deserve mention.

One such consideration is raised by Sterelny and Fraser themselves.<sup>71</sup> Norms are not only tools of cooperation—they are tools of *coordination*. Norms establish shared expectations for behavior. These expectations are internalized, and deviations are heavily punished. This feature incentivizes conformity, but it also carries a danger, for punishment can stabilize destructive behaviors as well as cooperative ones.<sup>72</sup> Even when the status quo runs *against* profitable forms of cooperation, then, agents can still have strong incentives for compliance.

Norms are also tools of identification. Cultures differentiate themselves through “ethnic markers” such as patterns of speech, dress, and dietary preferences. Patterns of normative response are differentia as well: members of a culture dress, dine, *and* moralize like one another. Importantly, moralizing is not just a matter of paying lip service to social mores. Talk is cheap. (One need not be committed to the cause to denounce the freeloader who spent the afternoon slacking off.) Thus, groups often demand costly signals of commitment to their way of life. Though these costly displays can promote group cohesion, they do not always promote stable and profitable forms of cooperation. Many signals of religious commitment, for instance, impose nontrivial opportunity costs.<sup>73</sup>

These considerations caution against taking a norm’s existence as a facilitator of cooperation as a reliable sign that it forms part of a near-optimal

70 Barkhausen, “Reductionist Moral Realism and the Contingency of Moral Evolution,” 677.

71 Sterelny and Fraser, “Evolution and Moral Realism,” 1001.

72 Boyd and Richerson, “Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups”; Abbink et al., “Peer Punishment Promotes Enforcement of Bad Social Norms.”

73 See Bulbulia, “Religious Costs as Adaptations That Signal Altruistic Intention.”

normative package. Moral norms have a lot of work to do. They must support stable cooperative ventures. But they must coordinate and delineate as well, all while remaining sensitive to changes in the social-environmental landscape. Yet all of this merely explains why nonoptimal packages *arise*; it does not explain their stubbornness. Just why do suboptimal moral packages *persist*?

One part of the explanation for the entrenchment of nonoptimal packages is that norms are not modular.<sup>74</sup> Norms form part of interconnected webs of cultural standards and expectations; it can be hard to modify one element of that web without making drastic changes to the rest. Human groups also tend to be normatively homogenous.<sup>75</sup> This leaves less room for suboptimal norms to be selected against in favor of the superior ones on offer (there simply *are not* any superior ones on offer).

Moreover, escaping from suboptimal packages often requires solving difficult collective action problems. A useful illustrative example is the practice of female foot binding.<sup>76</sup> Initially, foot binding functioned as a high-cost signal of status: only the wealthy could afford to immobilize potential workers. But it lost this signatory value when it became universal practice. At this stage, *everyone* was worse off than they were without the practice. Yet unilateral defection was no longer an option in a world where unbound feet meant poor marital prospects. A similar lesson applies to inegalitarian social arrangements. Just about everyone is worse off when the distribution of the cooperative surplus is radically unequal. But unilateral revolt is not a viable strategy; a successful revolution requires a critical mass of dissenters.

Finally, humans engineer their social worlds. If you are at the top, then you will presumably want to stay there. A legitimizing ideology is a wise investment—power tends to have a longer shelf life once you have convinced others that you have the celestial tick of approval.<sup>77</sup> The elite need not be swindlers, to be sure. Over time, they may well come to believe their own propaganda. Moral conviction is sadly not immune to the influence of self-interest. It is not in the least bit surprising that the institution of slavery was favored by those who stood to gain economically, nor that the elite often have a penchant for social stratification.<sup>78</sup>

74 Sterelny, “SNAFUS,” 325; Buchanan and Powell, “De-moralization as Emancipation,” 121.

75 Sterelny, “SNAFUS,” 325.

76 See Sterelny, “SNAFUS.”

77 Sterelny, *The Evolved Apprentice*, 111.

78 See Enoch, *Taking Morality Seriously*, 192–93; Buchanan and Powell, “De-moralization as Emancipation,” 122–23. As I hope the discussion here makes plain, it is possible for a social arrangement to be “entrenched” without being “stable” (though I admit this sounds strange). To take a more familiar example: a business might be excessively micromanaging, bureaucratic, and procrustean in its handling of employees—and these arrangements may

So far, I have argued that we should be wary of inferring from the presence and perseverance of a norm as a facilitator of cooperation that it forms part of a near-optimal package. But one may still demand a positive rationale for thinking that dreadful norms *do not* form part of such packages. On the face of it, honor killings and slavery support cooperation rather well. It certainly is not *obvious* that societies that endorse these behaviors are less stable than those that do not. A fully responsible treatment of this issue would require consideration of normative packages on a case-by-case basis—a Herculean task that I cannot hope to undertake here. Let me, however, provide some principled grounds for thinking that norms that heavily discount the interests of large subsections of the population are unlikely to be near-optimal.

The first thing to note is that unfairness breeds resentment, and that resentment breeds instability in turn. Following Phillip Kitcher, the “technological possibilities for violent retaliation now increasingly available to the poor” mean that radically inegalitarian societies often have a strong potential for collapse.<sup>79</sup> Though certain forces can and do entrench unfair arrangements, it does not follow that such arrangements are robust to any changes. Indeed, many factors threaten to bring these systems crumbling down.

For one thing, oppressive norms can often be difficult for large subsections of the population to internalize. Elliot Turiel documents astounding resistance to sexist mores. Among these is the example of women in Saudi Arabia, who protested laws refusing them the right to drive by driving a convoy of cars through the city of Riyadh.<sup>80</sup> Kristen D. Neff also found that lower-middle-class Hindu women in India are highly critical of their lack of independence.<sup>81</sup> Gerry Mackie reports that many women in cultures that practice female circumcision strongly disapprove of it.<sup>82</sup>

Norms that are not fully internalized have a strong potential for disintegration. When agents do not value compliance for its own sake, they must be provided with strong incentives to play along—usually, heavy penalties for

---

very well come to be entrenched as the result of market forces. But insofar as such policies breed contempt, they are unlikely to be stable; employees will, for instance, likely engage in efforts to undermine the system through strikes or other forms of protest. Even if these bureaucratic arrangements come to be entrenched, then, they seem unlikely to be stable in the long run. At the very least, they seem far *less* stable (and thus, further from optimal) than salient alternative systems.

79 Kitcher, *The Ethical Project*, 311. See also Railton, “Moral Realism,” 191–94.

80 Turiel, *The Culture of Morality*, ch. 9.

81 Neff, “Judgments of Personal Autonomy and Interpersonal Responsibility in the Context of Indian Spousal Relationships.”

82 Mackie, “Female Genital Cutting,” 143.

noncompliance. Yet this introduces a danger: as soon as the penalties break down, so too does the compliance. Antifornication norms are a nice illustration.<sup>83</sup> Sexual impulses are strong motivators. Reining them in requires penalizing promiscuity, through stigmas attached to illegitimate children, for example. As soon as contraception and urbanization appeared on the scene, these disincentives were not nearly as powerful. (Contraception removes the threat of illegitimate children. Urbanization affords greater privacy.) Moreover, such mechanisms of policing *are themselves* a significant social cost. Norms that must be enforced on a sizable portion of the population are *ipso facto* expensive norms to have.

In summary, then, near-optimal packages require stability (among other things), and stability is difficult to achieve when the interests of large subsections of the population are heavily discounted. This basic insight goes a considerable way toward helping the variety of naturalism under investigation avoid countenancing perverse normative packages. Insofar as the norms that make up these packages are not plausibly viewed as near-optimal, the naturalist need not say that they reflect the moral truths. This is not to deny that naturalists still have their work cut out for them. Any particular recipe for identifying the moral among the natural will still need to contend with the full suite of available empirical data. A more thorough defense of the idea that perverse norms are (relatively) ineffective at promoting stable, efficient, cooperation would require getting into the weeds to a greater extent than I have here.

On that note—and in the spirit of inspiring further optimism—let me edge closer to a conclusion with a more concrete case study. According to Christopher Boehm's well-known work on hunter-gatherer groups, capital punishment is not an uncommon response to reciprocity violations.<sup>84</sup> One may want to claim that these excessive punishment norms are morally perverse. And yet, they seem to have proven effective in stabilizing hunter-gather societies and their broadly egalitarian social arrangements for quite some time. Is the naturalist not then forced to view these punishment norms as reflecting moral truths?

Everything here will, of course, depend upon the details. To begin with, we should not overstate the extent of capital punishment in response to reciprocity violations; fewer than half of the groups Boehm studied (24 of 50) reported it, and ostracism and shaming are far more frequent reactions than moralistic killing.<sup>85</sup> It is also important to consider *which sorts of* reciprocity violations are typically met with capital punishment. Among the most common are those

83 Buchanan and Powell, "De-moralization as Emancipation," 113–14.

84 Boehm, *Moral Origins*.

85 Boehm, *Moral Origins* and "The Moral Consequences of Social Selection," 172, 174.

that involve an individual intimidating other group members—for instance, through “psychotic aggression” or “repeated murder.” When the bully is only perceived as a moderate threat to group functioning, nonlethal measures are often used instead.<sup>86</sup> These details may lead us to question whether hunter-gatherer punishment norms are *obviously* perverse (in the manner that, say, genocide or slavery are—the permissibility of capital punishment for murder is, after all, still a matter of live debate). They might also lead us to question *which* capital punishment norms really do the work of stabilizing cooperation. Given its relative frequency, it may well be that capital punishment in response to repeated murder bears the bulk of the explanatory burden here—as opposed to say, capital punishment in response to theft or taboo violations. And this latter point, of course, feeds into a more general question: namely, whether these normative packages are reasonably viewed as *near-optimal*. As Boehm notes, capital punishment is incredibly costly in the hunter-gatherer context insofar as it cuts off reproductive opportunities and limits social and familial support networks.<sup>87</sup> Even if excessive punishment norms stabilize cooperation to some degree, then, it is not unlikely that cooperation could be rendered more stable and effective still under alternative, less excessive arrangements.

In general, empirical questions such as these will clearly remain important to any recipe for identifying the moral among the natural of the sort that has occupied my attention here—that is, to the sort of recipe that singles out moral properties by appealing to certain *nonmoral* facts, such as the norms that stabilize efficient cooperation, or how language users would apply moral terms following negotiation and reflection.<sup>88</sup> What I have sought to show is that this sort of recipe turns out to be far more promising than it initially appears; it is far from being a foregone conclusion that all implementations of it will send us plummeting headfirst into the realm of perverse moral norms. *Some* implementations might, of course. But this just seems like a reason for thinking that some naturalists have gotten their particular recipe wrong rather than an indication of a problem with having such a recipe.<sup>89</sup>

86 Boehm, *Moral Origins* and “The Moral Consequences of Social Selection,” 173.

87 Boehm, “The Moral Consequences of Social Selection,” 174.

88 Sterelny and Fraser, “Evolution and Moral Realism”; Jackson, *From Metaphysics to Ethics*.

89 For a point of comparison: a particular functionalist analysis of mental states may end up counting the wrong kinds of things as desires. But this would not necessarily be a problem with functionalism or its naturalist ambitions; it would instead be a problem with that particular way of using functionalism (or, more perspicaciously, with the particular background theory being put to use). As with just about any method or schema, what we get out depends upon what we put in.

Naturalists who do not share my optimism may prefer a different sort of recipe. The most salient alternative would be one that singles out moral properties by appealing to certain *moral facts*. One might tie moral truths to the judgments of a morally reasonable person, for example, or to judgments that are interpersonally justifiable.<sup>90</sup> One benefit of this approach is that it seems well suited to fending off perverse norms; perhaps we simply cannot hope to get the right moral results *out of* our naturalistic recipe without putting the right moral ingredients *in*. However, I am inclined to agree with Bart Streumer that this alternative approach faces insurmountable difficulties (most notably, a problem of vicious regress).<sup>91</sup> But naturalists who disagree will, I hope, still be able to extract an important lesson from this paper: we may be able to go much further with a purely descriptive recipe than has previously been thought.

#### 4. CONCLUSION

My organizing focus in this paper has been the naturalist's prospects for accommodating the intensional and extensional character of morality. My organizing ambition has been to build a case for an optimistic prognosis. I do not pretend that these are the *only* challenges that moral naturalism faces. One not-too-distant cousin of the extensional challenge, for instance, appeals to the unsettling *arbitrariness* that the naturalist seems content to tolerate. In the naturalist's way of seeing things, the only thing to recommend our own package of norms over other possible contenders seems to be that—owing to idiosyncratic features of our history and psychology—such norms promote profitable, stable forms of cooperation among us. Given this outlook, it is difficult to see what *justifies* our norms over alternatives that achieve the same ends for other possible versions of ourselves.

I cannot hope to offer a response to this additional challenge here. But my arguments do suggest a natural line of reply. Should we ever arrive at a near-optimal normative package, we will have arrived at a way of getting along that is well suited to the creatures that we are. Contrary to what initial appearances may suggest, the fact that this normative package will be well suited to us seems *far from* arbitrary, for its suitability will be explained by deep-seated and relatively inflexible features of our social existence and psychology. "It works for us" might

90 See Brink "Realism, Naturalism, and Moral Semantics," 175–76.

91 See Streumer, *Unbelievable Errors*, 55–57. Streumer is also skeptical about the prospects of the recipe that I favor, which he takes to fall prey to what he calls "the false guarantee objection." See Streumer, *Unbelievable Errors*, 47–55. This objection is similar to what I have called the extensional challenge—hence my disagreement with Streumer's grim assessment of it.

sound shallow. But it sounds far less shallow once we remind ourselves just what is required for a moral package to work for us: it must resonate with us, coordinate us, and promote profitable cooperative enterprise among us. As I have been concerned to emphasize, not just any mode of moral interaction fits this bill. Any that does will have to build upon the very features that make us human.<sup>92</sup>

*University of Leeds*  
j.m.isserow@leeds.ac.uk

#### REFERENCES

- Abbinck, Klaus, Lata Gangadharan, Toby Handfield, and John Thrasher. "Peer Punishment Promotes Enforcement of Bad Social Norms." *Nature Communications* 8, no. 1 (September 2017): 1–8.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association, 2013.
- Barkhausen, Max. "Reductionist Moral Realism and the Contingency of Moral Evolution." *Ethics* 126, no. 3 (April 2016): 662–89.
- Boehm, Christopher. *Moral Origins: The Evolution of Altruism, Virtue, and Shame*. New York: Basic Books, 2012.
- . "The Moral Consequences of Social Selection." *Behaviour* 151, no. 2–3, (November 2014): 167–83.
- Bosman, Ronald, and Frans Van Winden. "Emotional Hazard in a Power-to-Take Experiment." *The Economic Journal* 112, no. 476 (January 2002): 147–69.
- Boyd, Richard N. "How to Be a Moral Realist." *Contemporary Materialism: A Reader*, edited by Paul K. Moser and J.D. Trout, 307–70. New York: Routledge, 2005.
- Boyd, Robert, and Peter J. Richerson. "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups." *Ethology and*

92 I am very grateful for the feedback that I received on this paper at several workshops and conferences over the years. I thank audiences and prereaders at the Origins of Social Inequality, Hierarchy, and Social Complexity Workshop at the Australian National University; the Philosophy of Biology Workshop at Dolphin Beach; the Meta-Ethics Workshop at the Frankfurt Business School; and MetaEssen VI at the University of Duisburg-Essen. For comments on earlier drafts, special thanks are owed to Ellen Clarke, Edward Elliott, Guy Fletcher, Ben Fraser, Josef Holden, R. A. Rowland, Kim Sterelny, and Pekka Väyrynen. For comments on the latest version, I thank the two anonymous referees for this journal who took the time to provide constructive feedback. This work was supported by an Australian Research Council Grant for the project "The Origins of Inequality, Hierarchy, and Social Complexity" [grant number FL130100141].

- Sociobiology* 13, no. 3 (May 1992): 171–95.
- Brink, David O. *Moral Realism and the Foundation of Ethics*. Cambridge: Cambridge University Press, 1989.
- . “Realism, Naturalism, and Moral Semantics.” *Social Philosophy and Policy* 18 (Summer 2001): 154–76.
- . *Responsibility, Incompetence, and Psychopathy: The Lindley Lecture*. Lawrence: University of Kansas, 2013.
- Brosig, Jeannette. “Identifying Cooperative Behavior: Some Experimental Results in a Prisoner’s Dilemma Game.” *Journal of Economic Behavior and Organization* 47, no. 3 (March 2002): 275–90.
- Buchanan, Allen, and Russell Powell. “De-moralization as Emancipation: Liberty, Progress, and the Evolution of Invalid Moral Norms.” *Social Philosophy and Policy* 34, no. 2 (Winter 2017): 108–35.
- Bulbulia, Joseph. “Religious Costs as Adaptations That Signal Altruistic Intention.” *Evolution and Cognition* 10, no. 1 (2004): 19–38.
- Cavedon, Lawrence, and Sheila Glasbey. “Outline of an Information-Flow Model of Generics.” *Acta Linguistica Hungarica* 42, nos. 3/4 (1994): 227–45.
- Copp, David. *Morality, Normativity, and Society*. New York: Oxford University Press, 2001.
- . “Why Naturalism?” *Ethical Theory and Moral Practice* 6, no. 2 (June 2003): 179–200.
- Davidson, Donald. “Radical Interpretation.” *Dialectica* 27, nos. 3/4 (1973): 313–28.
- De Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. “The Neural Basis of Altruistic Punishment.” *Science* 305, no. 5688 (August 2004): 1254–258.
- Eisenberg, Nancy, Richard A. Fabes, Paul A. Miller, Jim Fultz, Rita Shell, Robin M. Mathy, and Ray R. Reno. “Relation of Sympathy and Personal Distress to Prosocial Behavior: A Multimethod Study.” *Journal of Personality and Social Psychology* 57, no. 1 (July 1989): 55–66.
- Enoch, David. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press, 2011.
- Findlay, Leanne C., Alberta Girardi, and Robert J. Coplan. “Links between Empathy, Social Behavior, and Social Understanding in Early Childhood.” *Early Childhood Research Quarterly* 21, no. 3 (2006): 347–59.
- Foot, Philippa. “Moral Arguments.” *Mind* 67, no. 268 (October 1958): 502–13.
- . “Morality as a System of Hypothetical Imperatives.” *The Philosophical Review* 81, no. 3 (July 1972): 305–16.
- . *Natural Goodness*. Oxford: Clarendon Press, 2003.
- Frank, Robert H. *Passions within Reason: The Strategic Role of the Emotions*. New

- York: W. W. Norton, 1988.
- Frank, Robert H., Thomas Gilovich, and Dennis T. Regan. "The Evolution of One-Shot Cooperation: An Experiment." *Ethology and Sociobiology* 14, no. 4 (July 1993): 247–56.
- Gauthier, David. *Morals by Agreement*. Oxford: Clarendon Press, 1986.
- Haidt, Jonathan. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. London: Penguin, 2012.
- Hamlin, J. Kiley, Karen Wynn, and Paul Bloom. "Social Evaluation by Preverbal Infants." *Nature* 450, no. 7169 (November 2007): 557–59.
- House, James S., Karl R. Landis, and Debra Umberson. "Social Relationships and Health." *Science* 241, no. 4865 (July 1988): 540–45.
- Hume, David. *Enquiry Concerning the Principles of Morals*. Edited by Tom Beauchamp. Oxford: Oxford University Press, 2005.
- Hurka, Thomas. *Perfectionism*. New York: Oxford University Press, 1993.
- Jackson, Frank. *From Metaphysics to Ethics*. Oxford: Oxford University Press, 1998.
- Korsgaard, Christine Marion. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.
- Kitcher, Phillip. *The Ethical Project*. Cambridge, MA: Harvard University Press, 2011.
- Kuhlmeier, Valerie, Karen Wynn, and Paul Bloom. "Attribution of Dispositional States by 12-Month-Olds." *Psychological Science* 14, no. 5 (September 2003): 402–8.
- Kurt, Fred, and Khyne U. Mar. "Neonate Mortality in Captive Asian Elephants (*Elephas maximus*)." *Zeitschrift für Säugertierkunde* 61, no. 3 (1996): 155–64.
- Kurth, Charlie. "What Do Our Critical Practices Say About the Nature of Morality?" *Philosophical Studies* 166, no. 1 (September 2013): 45–64.
- Liszkowski, Ulf, Malinda Carpenter, Tricia Striano, and Michael Tomasello. "12- and 18-Month-Olds Point to Provide Information for Others." *Journal of Cognition and Development* 7, no. 2 (November 2006): 173–87.
- Loeb, Don. "Gastronomic Realism—A Cautionary Tale." *Journal of Theoretical and Philosophical Psychology* 23, no. 1 (Spring 2003): 30–49.
- Mackie, Gerry. "Female Genital Cutting: A Harmless Practice?" *Medical Anthropology Quarterly* 17, no. 2 (June 2003): 135–58.
- Mameli, Matteo. "Meat Made Us Moral: A Hypothesis on the Nature and Evolution of Moral Judgment." *Biology and Philosophy* 28, no. 6 (September 2013): 903–31.
- Manne, Kate. "Internalism about Reasons: Sad but True?" *Philosophical Studies* 167, no. 1 (January 2014): 89–117.
- Neff, Kristin D. "Judgments of Personal Autonomy and Interpersonal

- Responsibility in the Context of Indian Spousal Relationships: An Examination of Young People's Reasoning in Mysore, India." *British Journal of Developmental Psychology* 19, no. 2 (June 2001): 233–57.
- Nichols, Shaun. *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford: Oxford University Press, 2004.
- Nickel, Bernhard. "Generics and the Ways of Normality." *Linguistics and Philosophy* 31, no. 6 (2008): 629–48.
- O'Connor, Cailin. "The Evolution of Guilt: A Model-Based Approach." *Philosophy of Science* 83, no. 5 (January 2016): 897–908.
- Parfit, Derek. *On What Matters*. Vol. 2. Oxford: Oxford University Press, 2011.
- Pradel, Julia, Harald A. Euler, and Detlef Fetchenhauer. "Spotting Altruistic Dictator Game Players and Mingling with Them: The Elective Assortation of Classmates." *Evolution and Human Behavior* 30, no. 2 (March 2009): 103–13.
- Railton, Peter. "Moral Realism." *The Philosophical Review* 95, no. 2 (April 1986): 163–207.
- . "Naturalism and Prescriptivity." *Social Philosophy and Policy* 7, no. 1 (Autumn 1989): 151–74.
- Shafer-Landau, Russ. "A Defence of Categorical Reasons." *Proceedings of the Aristotelian Society* 109, no. 1 (August 2009): 189–206.
- Southwood, Nicholas. *Contractualism and the Foundations of Morality*. Oxford: Oxford University Press, 2013.
- Sperber, Dan, and Nicolas Baumard. "Moral Reputation: An Evolutionary and Cognitive Perspective." *Mind and Language* 27, no. 5 (November 2012): 495–518.
- Stanford, P. Kyle. "The Difference between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation." *Behavioral and Brain Sciences* 41 (2018): e95.
- Sterelny, Kim. *The Evolved Apprentice*. Cambridge, MA: MIT Press, 2012.
- . "SNAFUS: An Evolutionary Perspective." *Biological Theory* 2, no. 3 (September 2007): 317–28.
- Sterelny, Kim, and Ben Fraser. "Evolution and Moral Realism." *British Journal for the Philosophy of Science* 68 (December 2017): 981–1006.
- Street, Sharon. "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues* 19 (2009): 273–98.
- Streumer, Bart. *Unbelievable Errors*. Oxford: Oxford University Press, 2017.
- Sturgeon, Nicholas L. "Moore on Ethical Naturalism." *Ethics* 113, no. 3 (April 2003): 528–56.
- Tangney, June Price, and Ronda L. Dearing. *Shame and Guilt*. New York:

- Guilford Press, 2003.
- Toi, Miho, and C. Daniel Batson. "More Evidence That Empathy Is a Source of Altruistic Motivation." *Journal of Personality and Social Psychology* 43, no. 2 (August 1982): 281–92.
- Tomasello, Michael. *A Natural History of Human Morality*. Cambridge, MA: Harvard University Press, 2016.
- . *Why We Cooperate*. Cambridge, MA: MIT Press, 2009.
- Turiel, Elliot. *The Culture of Morality: Social Development, Context, and Conflict*. Cambridge: Cambridge University Press, 2002.
- Vonasch, Andrew J., Tania Reynolds, Bo M. Winegard, and Roy F. Baumeister. "Death before Dishonor: Incurring Costs to Protect Moral Reputation." *Social Psychological and Personality Science* 9, no. 5 (July 2018): 604–13.
- Warneken, Felix, and Michael Tomasello. "Altruistic Helping in Human Infants and Young Chimpanzees." *Science* 311, no. 5765 (March 2006): 1301–3.
- Western, Bruce, Jeffrey R. Kling, and David F. Weiman. "The Labor Market Consequences of Incarceration." *Crime and Delinquency* 47, no. 3 (July 2001): 410–27.
- Woods, Jack. "Footing the Cost (of Normative Subjectivism)." In *Methodology and Moral Philosophy*, edited by Jussi Suikkanen and Antti Kauppinen, 166–190. London: Routledge, 2018.