https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

CONSTRUCTING NORMATIVE OBJECTIVITY IN ETHICS

By David B. Wong

I. The Problem of Moral Normativity

A prominent problem for all naturalistic theories of morality has been to account for the apparent normative force of a moral demand. By "normative force," I mean the scope of proper application of that demand, along with the conditions for its proper application. To characterize the normative force of a moral demand is to characterize to whom it properly applies and what conditions must be fulfilled for it to be properly applied. In this essay, the crucial issues about normative force are whether and in what manner a moral demand can properly apply to an agent regardless of whether that agent has inclinations or desires that would be served by conforming to the demand. In ordinary moral discourse, it seems that moral demands *can* apply regardless of the agent's possessing the relevant inclinations. To judge that U.S. officials have violated a moral duty by imprisoning foreign nationals without charges is not necessarily to refer to any inclination those officials have that would be satisfied by doing their duty. Indeed, one may judge so while assuming they have no such inclination. We talk as if the duty applies regardless of the existence of such a motivation.

That moral demands are generally regarded as inescapable or nonhypothetical in this sense has generally dictated two opposing responses: on the one side, attempts to validate the apparent inescapability of moral demands; on the other side, attempts to show that the appearance corresponds to a deep and pervasive error on the part of moral-language users. My response generally falls into the first category and is based on recognition of the role of morality and, more generally, the role of substantive practical reason in shaping basic human motivations. By "substantive practical reason," I mean to suggest that the apparatus for reasoning about what to do includes not simply rules of inference for passing from premises to practical conclusions about what to do, but also an array of reasons for agents to act in certain ways, where these reasons are situational features that weigh in favor of agents' acting in certain ways. My conception of practical reason contrasts with the instrumentalist construal that has been dominant within naturalistic approaches to morality: the conception of practical reason as incapable of dictating ultimate ends but rather a formal faculty for guiding the transitions from basic, exogenous (relative to reason) motivations to nonbasic motivations and ultimately to

DOI: 10.1017/S0265052508080096
© 2008 Social Philosophy & Policy Foundation. Printed in the USA.

actions. On the instrumentalist conception, normative demands on an agent ultimately derive from some motivation the agent already has.

There is a connection between the way that the apparent inescapability of moral demands poses an explanatory problem for naturalistic approaches to morality and the dominant construal of practical reason. I subscribe to a naturalistic approach distinguished by three features: first, it seeks to understand the role of morality in human life by drawing on the human sciences; second, it rejects any *a priori* method for yielding substantive truths shielded from questioning informed by empirical evidence; and third, it refuses to take moral judgments, even ones we are strongly inclined to believe as true, as corresponding to some *sui generis* portion of the universe, existing independently of human needs, feelings, and will. Right away, problems arise from such a naturalistic perspective when we look at the third feature alongside the apparent inescapability of moral demands. What kind of normative content can apply to agents regardless of their inclinations if there is no moral reality existing independently of such inclinations? The dominant instrumentalist conception of practical reason would seem to eliminate the possibility that even if there were any such normative content it could *necessarily* provide *everyone* reasons to behave morally. Rather, the existence of reasons *contingently* depends on the particular inclinations and will of the agent.

A classic example of the conclusion a naturalist might draw from this dilemma is Philippa Foot's 1972 essay "Morality as a System of Hypothetical Imperatives." She accepts that whether one has reasons to conform to a moral demand depends on whether one has the appropriate desires or inclinations that would be served by conforming to the demand. On Foot's instrumentalist interpretation of moral reasons, Immanuel Kant is simply wrong to hold that, regardless of the content of one's desires, it is irrational to disregard one's moral duty. It is true, she grants, that moral demands are nonhypothetical in applying to people whether or not they have any desires that would be served by being moral. Foot, however, has another way of explaining how moral demands are nonhypothetical. They are nonhypothetical in the same way that the demands of etiquette are.[1] The demands of polite behavior apply even to those who have no interests served by being polite. Rude behavior remains subject to criticism from the standpoint of the institution of etiquette. To say of someone that he is rude does not mean, according to Foot, that he has a reason to be polite or that he is irrational to be rude.

The same is to be said of morality, except that morality is taught with more emphasis. In general, it matters more to others that one should be

[1] Philippa Foot, "Morality as a System of Hypothetical Imperatives," *Philosophical Review* 81 (1972): 305–16. In her later work, however, Foot takes a very different position on the force of moral demands. See her "Does Moral Subjectivism Rest on a Mistake?" *Oxford Journal of Legal Studies* 15 (1995): 1–14; and her *Natural Goodness* (Oxford: Clarendon Press, 2001), chap. 4.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

moral than that one should be polite, and this accounts for the feeling of inescapability we generally attach to moral demands. Kant, in Foot's view, wrongly infers from this kind of inescapability that the moral law is a universally and necessarily binding law of practical reason. We can see the error in this kind of inference by recognizing that, in terms of scope and conditions for application, etiquette possesses the same kind of inescapability as morality. But no one supposes that there is a law of etiquette that is a binding law of practical reason. Thus, Foot preserves a kind of inescapability for moral demands by divorcing it from the having of reasons to conform to them.

Richard Joyce, in contrast, argues that moral judgments *purport* to have a normative grip directly on the agent. To grant that a moral demand applies to an agent and yet to deny that she necessarily has a reason to satisfy it sounds very odd. This is to deprive moral demands of the practical clout or "oomph" they purport to have, to invoke Joyce's felicitous technical term.[2] Joyce correctly observes that Foot could have reformulated her position to admit the existence of reasons to be moral that do not depend for their existence on the agent's having the right kind of desires. These are reasons offered by morality as an institution, such that, given its rules, for example, everyone has a reason not to harm innocent people. However, Joyce argues that this kind of reason does not supply the practical clout that we seem to attribute to moral demands in everyday moral discourse.

Consider the parallel case of etiquette. Everyone has reasons to observe the rules of polite behavior in the sense that the rules of the institution of etiquette require the various forms of polite behavior. Someone who cares not at all for the institution of etiquette can disregard the reasons it offers to wield one's dinner knife in the polite manner. The fact that reasons of etiquette lack practical clout may be an acceptable result, but it is not acceptable in the case of morality because we seem to expect moral reasons to have this clout.[3]

Moral reasons *seem* to be genuine in an institution-transcending sense, Joyce observes, such that they seem to yield genuine deliberative considerations for those to whom they are ascribed. But in the end, Joyce suggests, the appearance may be an illusion. Since Joyce does not see how nonhypothetical reasons could transcend institutions and the motivations of agents to whom they are ascribed (and I think it is fair to say that his version of a naturalistic perspective has something to do with this), he is prepared to attribute a fundamental error to moral-language users. In making moral demands on each other, we presuppose a kind of transcendent normativity that does not exist: there are no nonhypothetical reasons to be moral that are external to or transcend those specified by human

[2] Richard Joyce, *The Evolution of Morality* (Cambridge, MA: MIT Press, 2006), 202–3.
[3] Ibid., 194.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

institutions and the existing motivations of agents. It should be noted, however, that Joyce's version of error theory, unlike J. L. Mackie's version, does not end with a definite dismissal of practical clout, but rather with the expression of pessimism that it will ever be vindicated. Unlike Mackie, who flatly denies that moral reasons have the kind of practical clout that is attributed to them in everyday moral discourse, Joyce simply doubts that we will ever find a way of explaining how moral reasons could have such clout.[4] His position is closer to a kind of moral agnosticism, rather than being a straightforward embrace of Mackie's moral atheism.[5]

But why do we presuppose, in everyday moral discourse, that moral reasons do have practical clout? Joyce's explanation begins with the idea that moral beliefs promote helping, cooperative behavior and regulate interpersonal relationships in such a way as to avoid great harm as well as produce significant benefit for all concerned. He points to recent developments in evolutionary theory that could explain why helping and cooperative behaviors could have a genetic basis. Joyce's further hypothesis is that a moral sense, one of the functions of which is to produce guilt when one fails to be helpful or cooperative, could have an innate basis. This moral sense presupposes the concept of moral reasons that transcend institutions and apply independently of the individual's motivations. Moral conscience, conceived in this way, might provide the extra "oomph" we need to be helpful and cooperative even in circumstances where being so might seem to go against prudence. In the long run, being a steady cooperative partner could enhance one's reproductive fitness, suggests Joyce, expanding upon a theme taken from Robert Frank.[6] The practical clout of moral reasons, in other words, may be an adaptive illusion.

How compelling is Joyce's response to Foot? Joyce is correct, it seems to me, in viewing Foot's account of the inescapability of moral demands as deflating in its effect. On Foot's account, moral reasons could apply to all agents, but the sense in which they could apply seems purely formal. Yet there is something unsatisfying about Joyce's explanation of why we have a concept of transcendent moral reasons. According to him, the mistaken belief that such reasons have a normative grip on us independently of our existing motivations serves the function of nudging us

[4] Ibid., 223.

[5] J. L. Mackie, in *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin, 1977), holds that it is a presupposition of moral discourse that objective values exist. However, such values cannot exist, he argues, because they are supposed to have a "categorically imperative element" that is "objectively valid" (ibid., 29). He means that objective values are supposed to give anyone reasons to behave in certain ways—reasons that are not contingent on having inclinations that would be served by so behaving. Therefore, Mackie's position is that moral demands purport to give us noncontingent reasons to behave in certain ways, but such reasons cannot exist. Joyce's position is more that no one has given a plausible way of explaining how such reasons could exist. He stops short of declaring that there is no such way.

[6] Joyce, *The Evolution of Morality,* 121. See also Robert Frank, *Passions within Reason: The Strategic Role of the Emotions* (New York: Norton, 1988).

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

toward cooperation when prudence might move us the other way. But *why* should the belief in transcendent reasons have such a motivating effect on us? Joyce says that guilt over the thought of acting against such reasons might nudge us toward acting according to them, but *why* feel guilt if we can make no sense of their normative grip? Either the belief itself is sufficient to motivate us, which does not fit very well with Joyce's naturalism,[7] or it is simply the guilt that motivates us, but then the motivating effect of guilt suggests there is some prior motivation to act on behalf of others in the first place.

On two crucial points, Joyce is right. First, we do treat moral demands as if they are not contingent on their serving the agent's existing desires (on this, he and I agree with Foot). Second, we do treat moral demands as if they provide reasons for action (on this, he and I disagree with Foot). It sounds odd to say that Joe morally ought to stop cheating on his taxes and that he has no reason to do so. It sounds downright contradictory to my ears to say that Joe ought to stop cheating and that he has no *moral* reason to do so. Ought-to-do's and duties imply reasons for those who ought to and who have duties.

The next crucial question is, "What kind of reason is implied by moral ought-to-do's and duties?" The reason is clearly a justifying reason. Within the moral perspective, to say that Joe has a moral reason to stop cheating on his taxes is to say that there is something about his situation that justifies his ceasing to cheat: for example, the fact that he would be benefiting from services funded by other taxpayers but not contributing to these services. Perhaps this is what Joyce means when he says that everyone is supplied reasons by the institution of morality. Joyce thinks that something more than a justifying reason from the moral perspective is needed for moral demands to have the practical clout ordinarily attributed to them. Otherwise, a reason that justifies an action from the point of view of etiquette could have as much of a normative grip on the agent as a moral reason. As to the nature of what more is needed, Joyce seems to have in mind something that could license saying to the agent, "You

[7] Naturalists generally tend to accept David Hume's thesis on the motivational inertness of reason: a belief alone cannot move an agent; it can do so only in conjunction with some motivation that has the structure of desire. On Elizabeth Anscombe's way of distinguishing desire from belief, according to which desire aims to make the world fit with it rather than aiming to fit the world, see Anscombe, *Intention* (Oxford: Blackwell, 1957). I think the motivational inertness thesis should not be an *a priori* feature of naturalism, but rather should be accepted on the grounds that it is part of more fruitful explanations of human action. Antonio Damasio's work on what goes wrong with the practical reason of certain brain-damaged patients is a start at vindicating the empirical fruitfulness of Hume's thesis. Damasio's patients are distinguished by the normality of their theoretical reasoning, even about practical matters, and by the abnormal inability to act on reasoning, even in matters of prudential reasoning about which they are able to arrive at sound judgments in the abstract. The damage is to the parts of their brains that process emotions, and Damasio's thesis is that practical reasoning, in order to be effective, needs to have various action scenarios marked by bodily processes that correspond to emotion. See Antonio Damasio, *Descartes' Error* (New York: Avon Books, 1994), 50.

are making a mistake if you don't take this reason into account in your practical deliberation." And the kind of mistake is not merely going against the norms of an institution, whether the institution is that of etiquette or morality.

The intuition behind Joyce's objection to Foot is that reasons addressed to an agent, if they have legitimate normative force, must go deeper than being based in an institution that claims some kind of authority over the agent. They must go deeper, I think, in three ways. First, the normative force is conceived as rooted in something that is not something purely external to the individual. It is not just, "*They* judge me to be mistaken if I do not conform to this demand." Second, the root is not the same as prudential normativity. The normative force is not derived, for example, from concern about being punished by disapproving others. Third, there is a practical point to whatever normative force moral reasons possess, and a practical point that concerns the agent. Otherwise it's just "browbeating," to use Bernard Williams's term.[8] That is, it must be possible for the agent *to act* on the basis of recognizing them as moral reasons.

We are now in a better position to understand why Joyce thinks he has to resort to error theory. Moral reasons purport to have a deeper grip on the agent than seems explicable on the basis of an instrumentalist theory of reasoning. It seems as if we do believe in transcendent, nonhypothetical reasons that apply to the individual externally to his motivational system, yet the instrumentalist theory, on a naturalistic view, seems to be the clearest theory we have of how reasons have a grip on us. We must be erring in conceiving moral reasons to be nonhypothetical, and Joyce is looking for an explanation of why that error is committed. He suggests that the error has a use, and if it is not instrumental to the individual's motivation, perhaps it is instrumental to the "purposes" of natural selection.

This explanation does not work well, as I have argued. I will offer an alternative explanation that affirms the existence of reasons to be moral that can get a normative grip on the individual's motivations, even when the reasons cannot be explained in terms of what serves the individual's desires and inclinations. I will argue that there is a good naturalistic explanation for their existence and of how they could come to motivate individuals whose existing interests are not served by acting according to them. To see our way clear to such an explanation, we must let go of one assumption that has been dominating the discussion among philosophers such as Foot and Joyce: the assumption that the apparatus of practical rationality exercises normative force only on the individual who is already constituted as an agent and an individual self. In contrast to this assumption, I would argue that the learning of moral reasons must be recognized

[8] Bernard Williams says this about the idea of applying a reason requiring an individual to act even though that reason is "external" to the individual's subjective motivational set. See Bernard Williams, "Internal and External Reasons," in Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981), 101–13.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

as going into the very constitution of the agency and selfhood of the person. That is, moral reasons go deeper by influencing the very structure of motivation and by enabling practical deliberation itself. This might sound constructivist and Kantian. To the first, I plead a hearty "guilty." To the second, I plead guilty if there can be a naturalistic and relativistic Kantianism. So that might be "not guilty."

Section II lays the groundwork for an explanation of the existence of nonhypothetical reasons to be moral by explaining how moral norms play a crucial role in making cooperation possible by shaping and reinforcing the diverse and potentially conflicting array of human motivations so that they are better suited for cooperative life. Section III sets out the way that moral norms are related to moral reasons, and also sets out how a particular set of moral reasons gets established as part of the morality of a group. Section IV explains how moral reasons so characterized help to shape motivation by becoming "embedded" in motivational propensities. Section V argues that given the role that moral reasons play in shaping human motivation and guiding practical deliberation, moral reasons go into the construction of persons, that is, into the construction of who they are as persons. This last theme leads to a naturalistic explanation of how moral reasons can have a deeper normative grip on the individual than can be explained by an instrumental conception of practical reason. Moral reasons do not simply answer to the desires and inclinations that make us who we are; they help to shape and are embedded in the desires and inclinations that make us who we are, and they can be transcendent reasons in that sense. The dominance of the instrumental conception of reason makes it harder to see that moral reasons are not simply based on independently existing desires and inclinations. Section VI attempts to undermine this dominance by arguing that the concept of a self that extends over time is a concept constructed to meet the demands of social cooperation. Prudential reasons are reasons to act on behalf of the self as it extends into the future and are often taken to constitute the paradigm of reasons that are instrumentally based on desires and inclinations of the individual. But such reasons, like moral reasons, are constructed to promote human cooperation and go into shaping the individual's motivations, and are not merely based on the individual's motivations.

II. Some Scene Setting: The Cooperative Nature of Human Life and the Role of Moral Norms in Making This Possible

Let me begin by sketching a picture of the diverse array of motivational propensities that are plausibly considered to be innate to human beings.[9]

[9] Most but not all of the rest of this section draws from ideas expressed in chapter 2 of my *Natural Moralities: A Defense of Pluralistic Relativism* (New York: Oxford University Press, 2006).

Alongside their instincts for self-preservation, members of the human species (or many of them)[10] developed capacities of care for kin, a willingness to engage in mutually beneficial practices of cooperation with others if they show a similar willingness, a willingness to punish those who violate the agreements and norms that make cooperative practices possible (even when the expenditure of resources to punish cannot be justified on the grounds of pure self-interest), and some degree of altruistic concern for nonrelated others. Human beings developed all these capacities because they were fitness enhancing in an inclusive sense, a conclusion that much of the latest work in evolutionary theory supports.[11] While human beings certainly evolved as strongly self-interested creatures, they also evolved motivational capacities that allowed them to form cooperative bonds with each other. Human beings, then, were selected to have a diverse array of innate psychological tendencies that can potentially come into conflict with each other if they do not have ways of regulating and tempering the expression of these tendencies. Human nature can be profoundly ambivalent in this way.

Cultural norms, I suggest, play a large part in this structuring of motivation. The human capacity to regulate the self through cultural norms coevolved with biological traits. The long period of the Pleistocene during which human beings evolved social instincts overlapped considerably with the period in which people began living in social groups with cultural institutions. If culture was a partner in this biological evolution, then it is plausible to hypothesize that some of our biological traits, as anthropologists Robert Boyd and Peter Richerson have suggested, might prepare us to regulate ourselves through culture: for example, the disposition to follow the majority or to emulate the most successful members of one's group. (The latter strategy requires people to be selective imitators and to have at least an inkling of what good solutions to common problems are.)[12] Such traits could have conferred an evolutionary advantage on

[10] The reason for this qualification will surface later in this essay.

[11] On the selection mechanism for altruism toward kin, see W. D. Hamilton, "The Genetical Evolution of Social Behavior," *Journal of Theoretical Biology* 7 (1964): 1–52. On the mechanism for selecting a willingness to cooperate with others if they show willingness to cooperate, see Robert Trivers, "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology* 46 (1971): 35–56. For a theory of "group selection" as the mechanism behind concern for non-kin, see Elliott Sober and David Sloan Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press, 1998); and for another theory emphasizing the role of sexual selection in altruism, see Geoffrey Miller, *The Mating Mind* (New York: Anchor Books, 2000). For evidence supporting the existence of non-self-interested willingness to punish, and to reward others who cooperate, see Herbert Gintis, *Game Theory Evolving* (Princeton, NJ: Princeton University Press, 2000).

[12] Robert Boyd and Peter J. Richerson, *Culture and the Evolutionary Process* (Chicago: University of Chicago Press, 1985); Boyd and Richerson, *Not by Genes Alone: How Culture Transformed Human Evolution* (Chicago: University of Chicago Press, 2005). In the latter book, they define "culture" as "information capable of affecting individuals' behavior that they acquire from other members of their species through teaching, imitation, and other forms of social transmission." By "information," they mean "any kind of mental state, conscious or

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

members of a group by enabling them to adopt satisfactory solutions to problems that were worked out by other members. Individuals do not need to "reinvent the wheel" on their own but can instead follow cultural norms established over some period of time, provided that the group's environment changes slowly enough so that the solutions embodied in those norms remain satisfactory.

Moral norms, on my view, culturally evolved to promote beneficial social cooperation, not simply through requiring behavior that is cooperative and considerate of the interests of others, but also through encouraging, strengthening, and directing the sorts of feelings and desires that make people promising partners in social cooperation.[13] A virtue of this functional conception of moral norms is that it helps to organize and systematize many of the most central moral beliefs that appear across cultures and historical periods: beliefs that specify the conditions for permissibly killing or conducting aggression against other human beings; beliefs about the right to assign and distribute the basic resources needed to sustain life; and beliefs that require reciprocation of good for good. There is a lot of variation in how these beliefs are filled in with specific content and in the nature of the particular restrictions and distributions, but a common end these beliefs serve is the regulation and promotion of social cooperation.

Some prominent functionalist accounts of morality claim that the primary purpose for which morality is invented is to counteract the destructive effects of self-interest (e.g., Thomas Hobbes) or the limitations on our sympathies for others (e.g., G. J. Warnock, J. L. Mackie).[14] I favor a more complex functional picture, under which the profound ambivalence of human nature is managed in a variety of ways and not just through constraint of potentially destructive self-interest. Moral norms need to take into account the strength of self-interest in order to accommodate

not, that is acquired or modified by social learning, and affects behavior" (*Not by Genes Alone,* 5). I agree with Dan Sperber and Nicholas Claudière ("Defining and Explaining Culture," *Biology and Philosophy* 22 [2007]) in their observation that Boyd and Richerson's use of "information" is too broad in one sense—culture is better taken to include only *widely distributed* information—and too narrow in another sense—the relevant kind of information can be implemented not only in the form of mental representations but also in the form of behaviors, artifacts, and institutions.

[13] This does not mean that directly facilitating social cooperation is the only function of morality. Some moral norms take the form of character ideals and conceptions of the good life specifying what is worthwhile for the individual to become and to pursue. This intrapersonal function of morality comprehends what has been called the "ethical," as opposed to what might be called the "narrowly moral." Morality in the broader sense used here comprehends the ethical. This part of morality helps human beings to structure their lives together in a larger sense, i.e., not just for the sake of coordination with each other but also for the sake of coordination within themselves.

[14] See Thomas Hobbes, *Leviathan*, part 2, chaps. 13–17; G. J. Warnock, *The Object of Morality* (London: Methuen, 1971); and J. L. Mackie, *Ethics: Inventing Right and Wrong* (London: Penguin, 1977), chap. 5.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

that motivation and to encourage its *integration* with motivations that more directly lead to acting on behalf of others. Effective moralities, then, do not merely restrain actions from self-interest or encourage the development of opposing motivations, though they do these things. Effective moralities provide outlets for the expression of self-interest that can be consistent with the expression of other-regarding motivations. Self-interested motivation can clearly have undermining effects on social cooperation when it motivates noncooperation and aggression against others. However, in the right circumstances self-interest can support, rather than oppose, other-regarding motivations.[15] Arrangements that generate some self-interested return to other-regarding behavior can create an "ecological niche," Jane Mansbridge suggests, that helps to sustain that behavior. By making that behavior less costly, these arrangements can increase the degree to which individuals feel they can afford to indulge their concerns for others.[16] Rather than saying that an effective morality should always constrain self-concern and reinforce other-concern, I would suggest that it should often attempt to accomplish a productive balance or reconciliation between those types of concern.

Norms of reciprocity that require a return of good for good received play a crucial role in such reconciliation. The need to reconcile self- and other-concern appears first in family relationships. Across widely different cultures, there are duties to respect and to honor parents and others whose roles involve raising and nurturing the young. Performance of such duties constitutes a kind of return of good for good, though what is returned, of course, is not always the same kind of good as what was originally given. Sometimes the return is similar to the original good, as in the case of children's care of aged parents. Most other times, however, the return is a good that is fitting to the nature of one's relationship to those who have cared and nurtured: obedience and receptiveness to what is taught, for example. Perfectly selfless parents might not need such reinforcement, but profoundly ambivalent beings might not be able to do without it. Note, however, that the benefit of being reciprocated for an act of helping need not be greater than the cost of helping in order to reinforce that initial act of helping. It need merely generate enough return so that helpers feel they can afford to indulge whatever other-directed concerns they have.

Furthermore, we humans cooperate on a scale that is far larger than kin groups and the small groups in which it is feasible to directly reciprocate

[15] See Gintis, *Game Theory Evolving*.

[16] Mansbridge has suggested that while other-regarding motivations such as those stemming from empathy do exist in most individuals, they do not have infinite value. Herbert Gintis's portrait of *Homo reciprocans* suggests the same qualification. If the costs of benefiting others are very high, many will simply decline to pay. See Jane Mansbridge, "On the Relation of Altruism and Self-Interest," in Mansbridge, *Beyond Self-Interest* (Chicago: University of Chicago, 1990), 133–43.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

to those who have conferred benefits upon us.[17] Moral norms play a key part in the shaping of motivations appropriate to such widely ranging cooperation. Moral norms, for example, might encourage the favorable evaluation of those who engage in helping and cooperative behavior, regardless of whether one has directly benefited from that behavior. In what has been called "indirect reciprocity," others become favorably disposed through reputation toward the helper as a potential cooperative partner.[18] A morality can help this kind of reinforcement via indirect reciprocity through placing value on trustworthiness. Learning who is trustworthy is learning to assess others for their reliability as potential cooperative partners, and communication of judgments of trustworthiness helps to sort the cooperators from those free-riders who move from group to group to avoid identification and punishment. Gossip may have a moral function after all, and the negative aspect of that function might be at least as important as the positive aspect. There is evidence that cooperators tend to punish free-riders by ceasing to cooperate with them even if noncooperation costs more to the cooperators than cooperation with free-riders.[19]

Moral norms direct the manner of expression of self- and other-concern so that they are more compatible. Such norms reinforce and strengthen other-concern. The performance of these functions is crucial for making a go of human cooperative life. In the next section, I explain how moral norms are related to moral reasons.

III. What Moral Reasons Are: Their Relation to Moral Norms and Their Place in Moralities

Reasons are those considerations weighing in favor of or against an agent's doing something.[20] They are structured as three-place relations

[17] Explanations of altruism toward non-kin that rely solely on the idea of natural selection over genetic variations rely on rather special conditions being set in place. For example, Charles Darwin (in *The Descent of Man*) thought that natural selection sometimes operates on groups and not just on individuals, so that in the case of human beings, a tribe with members willing to sacrifice for other members will prevail in competition with other tribes with no such members, or will do well in adverse natural circumstances, and will therefore gradually predominate among the human species. This explanation, however, depends on groups' preserving the genetic differences between them as the ones with greater proportions of altruists win out over the ones with lesser proportions. Members of a winning group who intermarried with members of a losing group would undermine this special condition, but this seems to have been a common occurrence. In contrast, the tendency toward conformity with cultural norms might preserve intergroup variation and allow groups with prosocial cultural norms to win out over others and continue adherence to those norms even in the face of intermarriage with members of other groups. See Peter J. Richerson, Robert T. Boyd, and Joseph Henrich, "Cultural Evolution of Human Cooperation," in *Genetic and Cultural Evolution of Cooperation*, ed. Peter Hammerstein (Cambridge, MA: MIT Press, 2003), 368–69.

[18] Richard Alexander, *The Biology of Moral Systems* (New York: Aldine de Gruyter, 1987).

[19] See Gintis, *Game Theory Evolving*. Robert Trivers, in "The Evolution of Reciprocal Altruism," identified a crucial role for "moralistic aggression" (negative reactions to perceived violations of reciprocity) in helping to reduce the incidence of free-riding. However, it is Gintis who correctly points out that in many instances there is an altruistic element to the willingness to retaliate against free-riders.

[20] This section summarizes ideas expressed in chapter 2 of my *Natural Moralities.*

between an agent A, an action X, and a feature F in the agent's situation that weighs in favor of A's doing X. For example, A may have a reason to help B in virtue of B's being in imminent danger of being harmed and A's being able to help with no risk and low personal cost to herself. As I indicated in Section I, a reason defined in this way is a *justifying* reason. F is the feature that purportedly justifies A's doing X. A justifying reason is not necessarily one that would motivate A to do X. It is not necessarily a motivating reason, but I will explain in Section IV how a justifying moral reason can become motivating.

We may think of reasons in general as providing the basic vocabulary for thinking and talking about what to do. Judgments as to what a person morally ought or ought not to do or as to what is morally right or wrong for that person to do can be construed as judgments as to what relevant moral reasons require for that person. To think about what we morally ought to do, we need to identify the moral reasons relevant in the given situation, and in cases where more than one situational feature is relevant and the different features weigh in favor of different and incompatible actions, we need to judge what the balance of reasons requires.

Moral norms guide our thinking about what reasons require. Sometimes they simply identify and articulate kinds of reasons, such as reasons to do what one has promised or agreed to do, or reasons to tell the truth in communicative contexts where there is a standard expectation of truth-telling. Sometimes norms provide guidance in thinking about what the balance of reasons requires, as in a situation where there are reasons against killing other human beings and reasons permitting one to save one's own life. The relevant norm permits killing in self-defense. Sometimes a norm marks the relative centrality of a value to a morality, such as the norm requiring respect for individual liberty and autonomy or the norm that emphasizes the realization of one's humanity in relationship to others.

The content of reasons (e.g., which situational features go into the identification of what we have moral reason to do) and of norms (e.g., how to prioritize in cases of conflicts between reasons) is constrained by morality's function of promoting and sustaining cooperation. That is, a genuine moral reason is such that acting on it under the appropriate circumstances contributes to this function of morality. (The contribution that acting on any single reason might make to morality's functions might have to be assessed according to the ways in which it works with other reasons recognized within that morality.) Receiving the help of another person, for example, would constitute a reason to reciprocate. The moralities of particular societies yield a diverse and rich set of moral reasons that operate on much more specific levels (e.g., the forms of help that would create duties to reciprocate and the appropriate forms of reciprocation would be specified more concretely), but to have a chance at being genuine moral reasons, the reasons specified within these moralities must satisfy the general constraints.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

How does a set of specific moral reasons get established as part of the morality of a given group? I approach this question as a question about the concepts of moral reasons and how they acquire the reference they have.[21] Much recent work on the nature of concepts has undermined the "classical" model that posits necessary and sufficient conditions for their application. One of the proposed replacements for the classical model is prototype theory, according to which concepts include features possessed by their instances, the features embodying the average or most typical instances. To take a frequently used example in the prototype literature, the concept of dog includes features making up a kind of composite "everydog" (has four legs, a tail, emits barking sounds), and an object that is a candidate for falling under that concept is more likely to qualify the more it resembles the composite typical dog.[22] Another proposed replacement for the classical model is exemplar theory, which holds that having concepts involves the ability to call up particular instances that serve as the standards of comparison for candidate instances. Having the concept of a dog involves the ability to call up from memory particular dogs one has encountered, and one compares dog candidates to the closest exemplars to see if one gets a close enough match.

Concepts need not be limited to one structure. Indeed, some concepts might involve the acquisition of prototypes that are constructed on the basis of exemplars. A child might acquire her prototype of everydog on the basis of encounters with particular dogs she knew while growing up.[23] She might call up the dog prototype to categorize most dogs she encounters, but if she were to encounter a difficult case, she might recall an atypical dog exemplar that most closely resembles the present animal. Concepts of moral reasons seem to exhibit this kind of versatility. Consider the concept of a reason to help another person. In acquiring this concept, we might have had certain concrete situations identified as exemplars of a reason to help: a parent demonstrates for us what is to be done when a sibling falls down and hurts himself; the experience of an exchange

[21] Much recent discussion of truth and reference has employed the language of concepts where the language of meanings has previously held sway. I prefer the language of concepts because it has become the language for expressing and defending alternatives to the "classical" model which holds that a term (or the concept expressed by it) refers by virtue of a set of necessary and sufficient conditions embodied in the concept. Furthermore, much of the empirical evidence undermining the classical model has come from studies in cognitive and developmental psychology of ways that people categorize things, and these studies are couched in the language of concepts. See Eleanor Rosch and Carolyn B. Mervis, "Family Resemblances: Studies in the Internal Structure of Categories," *Cognitive Psychology* 7 (1975): 573–605.

[22] See Jesse Prinz, *Furnishing the Mind: Concepts and Their Perceptual Basis* (Cambridge, MA: Bradford Books, 2003), 51–72.

[23] See Andy Clark, "Connectionism, Moral Cognition, and Collaborative Problem Solving," in *Mind and Morals: Essays on Ethics and Cognitive Science,* ed. Larry May, Marilyn Friedman, and Andy Clark (Cambridge, MA: Bradford Books, 1998), 109–13. Jesse Prinz's "proxytype" model of concepts is supremely eclectic in incorporating prototype, exemplar, and other models of concepts. See Prinz, *Furnishing the Mind,* chap. 6.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

of gifts or favors is in many cultures an occasion exemplifying a reason to reciprocate. Alternatively, people can construct prototypes of such reasons that generalize over these exemplar cases, giving rise to a conceptual representative of a "typical" reason, say, to help or to reciprocate. On many occasions of classification, we might call up a prototype if the current situation seems typical, but in novel or borderline cases, we might call up from memory the closest exemplar and try to determine whether there is a close enough match. Peter Singer's argument that the relatively affluent have a strong duty to help famine victims drew much of its power from his analogy to the duty to save a drowning child if doing so merely required wading into a shallow pool and ruining one's clothes.[24] Saving the life of a child at very low cost to oneself may be an exemplar of what we have very strong reason to do (or very similar to such an exemplar), and Singer's argument is that helping famine victims at comparatively low cost to oneself is a very close match to the case of saving the child.

Recall that moral norms specify the situational features that go into the constitution of reasons (e.g., the situational feature of having received help is specified in a moral norm as a reason to reciprocate in some manner) and also specify the priorities among reasons in case these reasons require incompatible actions in a given situation (conflict may occur because a situation contains several morally relevant features that call for incompatible actions). People specify reasons and decide on priorities, and hence construct these norms, as they develop their cooperative life together and as they reflect on how well or how badly that life goes. The norms that emerge and get accepted within a group establish the truth conditions for moral judgments made by its members, since these judgments are equivalent to judgments about what there is reason to do or what the balance of reasons requires in a given situation. Therefore, the moral truth is constructed by human beings as they construct the norms that specify what there is moral reason to do, but the task of construction is subject to constraints on adequate moralities that spring from human nature and the functions of morality.

I have already implied one such constraint: a morality that lacked norms requiring reciprocation for help received, for example, would be a woefully inadequate morality, given the functions of morality and given the way human beings are. Another constraint follows from the way that morality works through a large degree of voluntary acceptance of its norms. If conformity to its norms and reasons depended solely on the threat of force or coercion, the costs would detract greatly from the benefits of cooperation. It makes sense that human beings evolved a system for regulating and promoting cooperation that governs in this way. A

[24] Peter Singer, "Famine, Affluence, and Morality," *Philosophy and Public Affairs* 1 (1972): 229–43.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

further step in the evolution of morality also makes sense for creatures who explain and justify their actions to one another: voluntary acceptance of moral norms came to be seen as based on their justifiability to those governed by them. Hence, another constraint on moralities is that justification for following the norms and reasons of an adequate morality cannot crucially depend on falsehoods. Norms that permit the subordination of the interests of some members to those of others often violate this constraint because they are rationalized on the basis of false presuppositions about the supposedly inferior capacities of the subordinated for rational thought and self-control.

I have identified a couple of the ways in which the set of norms that groups actually construct could fail to satisfy the constraints on moralities. A group could fail to include norms that are necessary for the promotion of cooperation. The justification the group gives to some norms could be false or inadequate. When members of a group judge what the balance of moral reasons requires in a given situation, they invoke the norms the group has constructed that specify what moral reasons there are and the kinds of priorities that obtain among them. They invoke these norms on the assumption that they are adequate and meet whatever constraints apply to them, and because this assumption can be wrong, there is always room for criticism and debate over the group's current norms and over the truth about what the balance of moral reasons requires.

In this section, I have specified the structure of moral reasons and the role of moral norms in a group's specification of what moral reasons there are and the priorities among them. I have given examples of constraints that apply to a set of moral norms that stem from the role of morality in promoting cooperation and from features of human nature. Morality performs this role through shaping human motivation. In the next section, I describe how moral reasons help to shape motivation.

IV. How Moral Reasons Help to Shape Motivation

The crucial concepts in my explanation of how moral reasons function in shaping motivation are those of motivational propensities, the intentional objects of such propensities, the embedding of moral reasons within the intentional objects, and the channeling of propensities through restructuring of intentional objects. *Motivational propensities* are functional states grounding dispositions to act or to feel under certain circumstances. They can take the form of a felt urge toward an *intentional object*, as in thirsting after a drink of cool water (where the object is the action of drinking the water), but it is not necessary for the intentional object to serve as a *phenomenological* content of a propositional attitude. That is, it need not be an object of awareness for the agent. When we specify the intentional object of a motivational propensity, we are specifying the motivational

direction of the relevant dispositions to act and the way they can form an intelligible bundle of action and feeling tendencies. Human motivational propensities are rooted in the imperatives of our biological being, but the basic directions of those propensities can be refined, rechanneled and reshaped. One of the ways in which these directions get changed is through our learning what reasons there are for doing things. The mere learning of such reasons does not assure any motivational propensity to act accordingly. The recognition of reasons must get *embedded* in the propensity.[25]

To support the possibility that the embedding of reasons can serve the purpose of channeling motivational propensities, let me point to the phenomenon of moods that become emotions. One way to distinguish moods from emotions is to say that moods are free-floating. They have no intentional object in the way that emotions have—no "aboutness." One is afraid *of* something, but one can be anxious about nothing in particular. Free-floating anxiety can easily turn into specific fears. Consider the case of a man who suffered from chronic feelings of anxiety. After the birth of his first child, he became frequently concerned about his child's safety, to the point of worrying that his child might one day climb onto the garage roof, fall off, and hurt himself on the stone bench below. This motivated the man eventually to hire workmen to break up the bench with sledgehammers and cart away the rubble.[26] This case concerns an excessive fear, even a pathological one (though some of us who are parents may admit to taking only slightly less ridiculous precautions for the sake of our children), but it dramatically illustrates how moods acquire intentional objects. Certainly, the man had reason to be concerned about the safety of his child: e.g., the child's vulnerability to harm and lack of awareness or judgment about what constituted dangerous play. This case also illustrates how reasons can get embedded in motivational propensities. That part of an emotion that is independent of the intentional object (the free-floating anxiety) can provide the motivational energy behind thoughts (the possibility of one's child falling onto a stone bench) that serve as reasons for action (having the bench broken up), and it thereby is channeled into that specific course of action.

The aforementioned case of the anxious man illustrates the mutability of our emotional lives. It has been remarked that flexibility of response is characteristic of creatures with emotions and that such flexibility con-

[25] The theme of reasons getting embedded in motivational propensities is expressed in chapter 7 of my *Natural Moralities,* and in my essay "Moral Reasons: Internal and External," *Philosophy and Phenomenological Research* 72 (2006). The characterization of embedding in the present essay, however, is revised from that previously published material, and is connected to empirical examples that I have not previously discussed. The application of this theme to the debate between the Footian and Joycean views of moral inescapability is not discussed in that previous work.

[26] This case is described by Gerald L. Clore and Andrew Ortony in "Cognition in Emotion: Always, Sometimes, or Never?" in *Cognitive Neuroscience of Emotion,* ed. Richard D. Lane and Lynn Nadel (Oxford: Oxford University Press, 2000), 27.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

ferred some evolutionary advantage on them. If you have a dog, as I do, who is very interested in chasing rabbits, you might have experienced the reaction of rabbits upon spotting a potential predator approaching from a distance. A common initial response is to freeze while blood flows to the muscles in preparation for fleeing. Whether the rabbit continues to freeze or runs away seems to depend in part on the rabbit's noticing whether the predator has zeroed in and is ready to pounce. Human beings have far greater flexibility of response in analogous situations. A member of a racial minority walks down the street of an unfamiliar part of town. As he notices the looks from those he is approaching, there might be a similar preparation for action, but a lot will depend on how he interprets the looks—as curious in a friendly way, as hostile, as getting ready to deliver an assaultive remark, and so on. As psychologists Gerald L. Clore and Andrew Ortony observe, "there seem to be two fingers on the emotional trigger: one controlled by early perceptual processes that identify stimuli with emotional value and activate preparation for action, and a second controlled by cognitive processes that verify the stimulus, situate it in its context, and appraise its value."[27] They observe that increased flexibility is connected to an expanded capacity for subjective experience, which "registers the urgency of a situation, provides information, and allows processing priorities to be revised." People "can entertain alternative courses of action and sample how they would feel about different outcomes," and in order to do this, "they must be aware of the stimulus that occasions the processing."[28] Thus, if our passerby registers looks of sullen but silent hostility, he may simply think to himself, "So what's new?" and walk on by without a further glance. He may consider throwing a hard stare back but may feel that the possible escalated outcome of such an action would not be worth his own emotional disturbance, much less the potential bodily harm. A further bit of complexity in this example is that our passerby can, in some sense, decide how angry he's going to get. He can decide how he will react to the situation and knows in advance that certain ways of reacting will escalate whatever feelings exist on the other side. He knows that he will react to such feelings, thereby raising the stakes on both sides.

Let me clarify. I am not claiming that emotions simply are cognitions of a certain sort but rather that cognition can enter into complex relationships with affective and conative components of emotions, producing new emotions, changing existing ones, and channeling the motivational directions of those emotions. The interactivity of cognition and emotion allows for an adaptive flexibility of response. This flexibility and capacity for a greater self-regulation of one's emotions characterizes human nature.

[27] Ibid., 41.
[28] Ibid.

It is not only a capacity exercised by individuals over themselves, but also a capacity exercised by groups over their members.

Here is an example of a use of moral reasons to shape motivation. In 1A7 of the *Mencius*, an account is given of a conversation between Mencius and King Xuan, the ruler of a Chinese state. Mencius is attempting to persuade the king to adopt the Confucian *dao* or way of ruling by becoming a true king to his people, namely a king who tends to their welfare rather than launching his state into wars of territorial expansion, thereby over-taxing and drafting his subjects into his army. The king wonders whether he really has the stuff to be this kind of king Mencius is describing, and Mencius replies that he thinks so, asking whether the following story he has heard about the king is true. The story is that the king saw an ox being led to slaughter for a ritual sacrifice. The king decided to spare the ox and substituted a lamb (that he did not see) for the ritual sacrifice. The king acknowledges that this story is true, and thinking back on that occasion, recalls that it was the look in the ox's eyes, like that of an innocent man being led to execution, that led him to substitute the lamb. Mencius then comments that this story demonstrates the king's capability to become a true king, and that all he has to do is to extend the sort of compassion he showed the ox to his own people.[29]

The conversation between Mencius and King Xuan portrays, in my view, an attempt by Mencius to embed a reason to care for one's people in the king's motivations. The king had shown a prior capacity for compassion in sparing the ox (even if he substituted the lamb, but that is another aspect of the story I cannot linger on), and Mencius is seizing on the king's recall of that past event in pressing for another expression of that capacity, but now directed toward his people. The king is no moral exemplar, but he is likely to have already learned that the vulnerability of his people to suffering (much of it potentially inflicted by him) is a justifying reason from the standpoint of morality for him to take certain actions and to refrain from others.

What would constitute that reason's being embedded in the motivational propensities that constitute the king's compassion? The king would see the potential suffering of his people as a reason to act, and that recognition would become associated with so acting, such that alleviating or preventing the suffering of his people would now be part of the intentional object of his compassionate motivational propensities. This does not mean that the king's recognition of his people's suffering would invariably result in the appropriate action, but if it did not, it would mean another set of appropriate reactions. One sort of reaction would be to question himself. He would question whether he had failed to do what he

[29] For a translation of this passage, see *Readings in Classical Chinese Philosophy*, ed. Philip J. Ivanhoe and Bryan Van Norden (New York: Seven Bridges Press, 2001), 116. The version of the text available today is dated from the second century C.E.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

had reason to do. Other appropriate reactions would be remorse and actions such as apology or an attempt to compensate those he had failed. Therefore, the embedding of a reason in motivational propensities means not only that the agent recognizes the reason, but also that recognition of the reason guides the exercise of the propensity in a way it had not before the embedding took place. And in cases where recognition of the reason fails to result in exercise of the propensity, embedding may still show itself in appropriate emotions (such as remorse) or actions (such as apology or attempted compensation).

How is such embedding accomplished? The answer to this must ultimately rely on a great deal of empirical study. Here are some ideas to test. Perhaps reminding the king of a previous exercise of his compassion, prompting him to recall it so that he relives and refeels that occasion, while reminding him that he has reason to alleviate or prevent the suffering of his people, is an effective way of trying to embed that reason in his compassion. Perhaps that is why dry philosophy lectures in normative ethics are so ineffective at bringing about genuine moral change in their audiences, and why films, narratives, and poetry can be more effective. The embedding of a reason is more likely if the propensity is activated at the time that the reason is being recognized by the agent. When the propensity involves complex emotion of the sort that can interact with new cognition, we have the possibility of transformation of emotion as the propensity is channeled.

Much of the most interesting work in the cognitive neuroscience of emotion and cognition is still very speculative. Allow me to speculate, however, on the neural mechanisms of embedding, using Antonio Damasio's theory of somatic markers.[30] Somatic markers are bodily reactions to a mental image and correspond to emotions. They ultimately derive from our biological drives, though they can become greatly modified through experience. A negative marker, attached to the mental image of a possible action and/or its outcome, can lead one immediately, before any cost-benefit analysis, to reject an option, roughly in the way one might recoil from the sight of a snake. Somatic markers can get connected to predicted future outcomes of certain scenarios. In deciding whether to alienate a friend for the sake of financial gain, one might picture the look on his face when he finds out what you've done, for example. Positive markers are beacons of incentives; that is, they highlight a certain response option. One of Damasio's hypotheses is that when different somatic markers are juxtaposed to different combinations of images, they modify the way the brain handles those images, and thus operate as a bias. The bias might allocate attentional enhancement differently to each component image, the consequence being the automated assigning of *varied degrees* of attention to *varied contents*.

[30] Damasio, *Descartes' Error,* 173–89. Damasio is a behavioral neurologist and neuroscientist at the University of Southern California.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

This hypothesis rings true to me phenomenologically and also as an explanation of how others attach greater weight to certain reasons rather than others. Damasio's picture provides a way of making sense of the operation of Aristotelian habituation, to which John McDowell refers when he talks about the acquisition of a "second nature" in the process of socialization.[31] Consider the example of the negative marking of the scenario of a friend's face when he finds out you've betrayed him for the sake of financial gain. The negative marking biases your attention toward that scenario in considering all the consequences of the betrayal, perhaps eclipsing alternative scenarios of what you would do with the money. It might be so strong a negative marker that it exhibits the phenomenon that McDowell mentions as characteristic of a virtuous person's deliberations: a moral reason "silencing" other nonmoral considerations that, in other contexts, might weigh significantly in practical deliberation. The thought of betraying a friend "silences" other considerations in altogether diverting your attention from the prospect of financial gain that normally weighs in favor of a course of action.

Damasio's somatic marker theory might be part of the explanation of what happens when moral reasons become embedded in motivational propensities. The situational features designated by reasons get somatic markers attached to them. If Mencius succeeds with King Xuan, the somatic marker that was attached to looking into the eyes of the terrified ox gets attached to images of his people suffering from the wars he embarks his state upon.

My suggestion in previous sections has been that human beings evolved to guide themselves through cultural norms that specify reasons and establish priorities among reasons. In this section, I have given an account of the role that moral reasons play in the shaping of motivations. In moral education and socialization, these reasons become embedded in motivational propensities and render individuals more suitable for the forms of cooperative activity that characterize human life. By shaping motivational propensities that can become central to who we are, reasons can enter into the kinds of persons we are.

V. The Challenge of Accounting for Normative Objectivity: How Reasons Go into the Construction of Motivations and Persons

Let me return to the problem of making sense of nonhypothetical moral requirements that are external to individuals' existing motivations. Such requirements purport to have a deeper grip on the agent than seems explicable on the basis of an instrumentalist theory of reasoning, according to which these requirements must concern actions that serve some existing

[31] John McDowell, *Mind and World* (Cambridge, MA: Harvard University Press, 1996), 84.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

motivations of the agent. Philippa Foot responds by construing the requirements as social demands that individuals only have reason to satisfy when it would serve some of their motivations to do so. Richard Joyce correctly sees that these demands purport to have a deeper normative grip on individuals, but he does not see how individuals necessarily would have reasons to satisfy these demands, independently of their existing motivations. He does not see this because he holds an instrumental theory of rationality.

I have proposed an alternative theory of rationality that includes not only forms of correct inference from beliefs to beliefs, and from beliefs and desires to desires and intentions, but also substantive conceptions of what people have reason to do. Such reasons are social constructions in the following sense: the explanation of why we have the moral reasons we have is that they play a crucial role in shaping our motivational systems for the cooperative life. The kinds of reasons we have are not just out there independently of this function they have in enabling human life. Furthermore, they are constructed in the sense that the specific form of the reasons we have is a variable matter that is not dictated by constraints on morality following from its functions and human biological traits. These functions and traits certainly do constrain how we attempt to shape motivation for the sake of cooperation, but they do not dictate specific reasons. For example, moralities are constrained to require reciprocation for the help of another, but a rich and diverse set of moral reasons specify the specific circumstances under which one has a reason to reciprocate and in what manner one should reciprocate.

Moreover, I hold that such functional constraints on morality underdetermine what a morality must look like in another sense. In a great number of cases, such constraints do not dictate how a morality guides people in cases of conflict between important values such as special duties to reciprocate to particular others such as members of one's family (on the one hand) and impersonal duties to care for and act on behalf of strangers (on the other hand). Nor do the constraints dictate whether a morality is to emphasize the value of belonging to and being responsible to a community that is defined by a set of relationships or the value of being left to choose for oneself the kind of life one is to live independently of the community to which one belongs and its social interests. Moralities distinguish themselves from one another precisely in how they guide people's actions in the case of value conflicts such as conflicts between special and impersonal duties or between community and autonomy. Guidance is provided by particular value priorities, and these priorities are constructed in the sense that they are not there independently of the ways of life that people construct for themselves.[32]

[32] This theme leads to the way in which my theory is relativistic: moralities that are incompatible in the sense of requiring different patterns of action, at least in some types of situations, may correspond to different sets of truth conditions established for moral judg-

Constructed reasons and constructed moralities contribute to constructed persons. Consider again the idea that human biological nature coevolved with the capacity to act on cultural norms. Moral reasons are an important subset of cultural norms and play an especially important role in fitting persons for the cooperative life. Moral reasons fit persons for such a life if they get embedded in motivational propensities: for example, propensities regarding how much persons care for others; which others they care for, and in what degree and in what ways; and whether the special duties of family take precedence over impersonal duties to strangers (and, if so, under what circumstances).

Of course, such embedding does not always take place, and there may be a number of causes for this result. One cause stems from the polymorphism of human motivational traits. For example, one of the hypothesized mechanisms by which altruistic traits could be selected is group selection, under which altruism is selected because it increases the fitness of some groups over other groups. Such a hypothesis does not require the "winning" groups to have nothing but altruists, but only a relatively high concentration of altruists compared to their competitors. Thrasymachus might be the spokesman for the purely self-interested cohort that group selection might have left in place. Members of this cohort are, no less than altruists, shaped by what reasons are constructed in their societies, but in their case, the most influential reasons get embedded in self-regarding propensities. Whether they behave morally might depend crucially on the willingness of others to invest their resources in punishing noncooperators. Another cause for the failed embedding of moral reasons is that cultures provide and shape individuals with different kinds of reasons—the prudential kind, for instance, as well as much more specific reasons, such as the beauty of an elegantly formed sentence as a reason for learning the craft of writing. Institutions and practices within a society may have their own subcultures and may stress reasons other than the moral kind. The degree of importance they put on such reasons can create tensions and outright conflicts with morality. As an increasing number of anthropologists have recently come to recognize, cultures are typically not harmonious and coherent but internally diverse with many different competing strains, containing corresponding differently motivated individuals and significant motivational ambivalence within individuals.[33]

ment. These different sets of truth conditions may refer to common types of moral reasons, but embody different value priorities in providing instructions as to how to balance and prioritize conflicts between different moral reasons when they apply to the same situation. See chapter 2 of my *Natural Moralities.*

[33] The anthropologist Bradd Shore gives, in a study of Samoa, a striking example of a culture's internal diversity and of the resulting ambivalence. Following the violent murder of his father, a young man received public counsel from a village pastor in formal Samoan that he must resist the temptation to avenge his father's death, and must keep in mind the values of peace and harmony and forgiveness. Yet, later, this same pastor, this time in colloquial Samoan, warned the young man that if he *failed* to kill the murderer of his father,

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

Consider that many of the most compelling identities that people possess are social-role identities. To have such identities is to be practically oriented so as to perform the responsibilities and exercise the rights of the relevant roles. In this regard, social roles provide their own sets of reasons for action—situational features to which people inhabiting these roles are prompted to respond in certain ways. Some social roles provide reasons that overlap with or are included within moralities; for example, being a parent provides reasons for responding to children's needs and nurturing their capabilities. However, the reasons provided by such moralized social roles might not always lead to moral action, as when parents zealously strive to meet their children's needs at the cost of neglecting and undercutting the welfare of other children. Other social roles, such as those created within economic institutions and practices, create even more potential conflict with moral reasons. Social ideals that are championed by economic institutions may compete with moral ideals as well as affect the content of these ideals.

Consider persons in whom moral reasons are not well-embedded or are embedded in motivational propensities that are relatively weak compared to others. Are such persons irrational for not being motivated by moral reasons? Here I agree with those who think that irrationality connotes a more blatant error in reasoning than failure to appreciate some of the reasons there are (where failure to appreciate means failure of embedding in motivational propensities). Irrationality connotes things such as the embrace of blatant inconsistencies.[34] Failure to appreciate moral reasons falls into the realm of the *unreasonable*, where this means something like not being a suitable cooperative partner, but fellows such as Thrasymachus would not be particularly disturbed by that label. Moral reasons, then, have a deep normative grip on the individual in being part of the apparatus of practical reason with which individuals are socialized and made into rational agents. But the depth of the grip does not make it such that every individual will be irrational in failing to be motivated by the normative force of *moral* reasons.

The conception of practical reason defended in this essay bears an important resemblance to the kind of expressivism defended most prominently by Allan Gibbard.[35] He denies, as I do too, that reason is purely instrumental. Talk about "what makes sense" to do and feel emerges from the efforts of human beings to coordinate with one another. Gibbard

he would not be his father's son. See Shore, "Human Ambivalence and the Structuring of Moral Values," *Ethos* 18 (1990): 165–79. For an influential critique of the older view of culture as a static, uniform whole, see Renato Rosaldo, *Culture and Truth* (Boston: Beacon Press, 1989).

[34] See John McDowell, "Are Moral Requirements Hypothetical Imperatives?" in McDowell, *Mind, Value, and Reality* (Cambridge, MA: Harvard University Press, 1998), 77–78; and Thomas Scanlon, *What We Owe to Each Other* (Cambridge, MA: Belknap Press, 1998), 25–26.

[35] Allan Gibbard, *Wise Choices, Apt Feelings: A Theory of Normative Judgment* (Oxford: Clarendon Press, 1992).

recognizes that talk about what makes sense or what we have reason to do and feel has the crucial function of shaping motivation for the sake of further cooperation (though he does not, to my knowledge, tell a story about reasons getting embedded in motivational propensities). He does not interpret normative talk as straightforwardly cognitive in content. I have articulated a view of how reasons-talk can perform an expressivist and shaping function while being straightforwardly truth-functional. Reasons are situational features that require of agents certain kinds of actions or omissions. Moral reasons are situational features identified in accordance with morality's function of promoting and sustaining social cooperation. Certain reason types will be universal, required by any morality that has a hope of performing its function. Others will be local, required by a particular morality but not necessarily by all that can perform morality's function. I have outlined a way that truth-functional talk of reasons can perform the function of shaping motivation through the embedding of moral reasons in the intentional objects of motivational propensities.

VI. Constructed Persons and Reasons of Prudence

The dominance of the instrumental conception of reason makes it harder to see that moral reasons are not simply based on independently existing desires and inclinations. This section attempts to undermine the instrumental conception's dominance by arguing that the concept of a self that extends over time is a concept constructed to meet the demands of social cooperation. Prudential reasons are reasons to act on behalf of the self as it extends into the future and are often taken to constitute the paradigm of reasons instrumentally based on desires and inclinations of the individual. But such reasons, like moral reasons, are constructed to promote human cooperation and go into shaping the individual's motivations, not merely answering to them.

To begin my case for this conclusion, consider Thomas Nagel's argument in *The Possibility of Altruism* that the interests of an agent's future self provide her with reasons to act *now*, even though no presently existing interest grounds such reasons.[36] Failure to recognize such reasons constitutes a failure in prudential rationality to recognize one's future self as equally real as one's present self. Nagel makes this argument in order to construct an analogy. The failure in prudential rationality is like the failure to recognize others as equally real as oneself. My focus here is not on the analogy but on Nagel's argument that it is a failure of prudential rationality to fail to recognize one's future self as equally real as one's present self. Nagel's argument is ultimately an appeal to our intuitions: Isn't it evident upon reflection that one's future self *is as equally real* as one's present self? I do not think this thought is so self-evidently true.

[36] Thomas Nagel, *The Possibility of Altruism* (Oxford: Clarendon Press, 1970), chap. 8.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

I agree that reasons of prudence include reasons to act now for the sake of interests we will have in the future, but such reasons are constructed for the sake of constructing agents who can integrate and balance among their interests over time. To learn that one has reasons to provide for the satisfaction of one's future interests is to learn such integration and balance. The question of what kind of balance is right to achieve is trickier than simply giving each of one's future selves equal consideration. One might think that all the stages of one's self are equally real and therefore deserve equal consideration, but the more distant those stages, the less certain one is about their existence, and moreover, the greater the likelihood of there being less psychological continuity and connectedness with those future stages. (Here I use "continuity" in the Parfitian sense to refer to memories of past experiences, intentions formed and then acted upon, resemblance over time between sets of beliefs, desires, and goals, and resemblance over time in character traits; and I use "connectedness" to refer to overlapping chains of psychological connections.)[37] My far-future self may be a much more admirable person than I am now, but for that reason, I may feel less connection with that person and, not unreasonably, may give that person less consideration than my near-future selves. In putting forward these considerations, I do not mean so much to undermine this particular conception of prudence as to indicate that the pragmatic purpose for having prudential reasons is consistent with significant variations in judgment about what we have prudential reasons to do. There is no *fact* such as the equal reality of various future selves I might be, from two minutes hence to twenty years down the line, that requires a particular conception of prudence.

Derek Parfit, in *Reasons and Persons*, has argued that, in effect, we may have no more reason to care about a future self than we do about other people. This is because there is no deep metaphysical fact that marks our separateness from other people. There would be such a fact if we were separate Cartesian thinking substances, but if one thinks, as Parfit does (and I do), that there are no such separate substances, the most plausible alternative view is that being one person over time is nothing beyond having a brain and body and a series of interrelated physical and mental events (where this relation is characterized in terms of psychological continuity and connectedness).

Christine Korsgaard has argued in response to Parfit that the requirements of practical agency in fact confer unity on the self. She agrees with Parfit that there is no deep metaphysical fact about being the same person over a stretch of time—no ontological entity or fact beyond having a brain and body and a series of interrelated physical and mental events (interrelated, that is, through connectedness and continuity). Rather, and here is where she departs from Parfit, the ground for thinking of oneself as

[37] Derek Parfit, *Reasons and Persons* (New York: Oxford University Press, 1984), 206–16.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

a unified agent is the raw necessity for eliminating conflict among our various motives and for the unity implicit in the standpoint from which one deliberates and chooses among these motives. On her view, this standpoint involves some principle or way of choosing that one regards as expressive of oneself and that provides reasons regulating one's choices among desires.[38] Furthermore, she observes that most of our projects extend over periods of time and that we think of our activities and pursuits as interconnected in various ways, comprising plans of life. The unity of the self is a consequence of the practical requirements for coherence over time and among motives and activities. On her view, then, the practical requirements of agency ground prudential reasons directly, and not by way of being founded on a prior notion of unity.

I accept Korsgaard's argument to this extent: the practical requirements of agency do contribute to the importance of identity over time, however we conceive of its basis. But the degree of practical coherence we need, and the duration over which we need it, seems a variable matter and subject to overriding by other practical concerns. A certain degree of conflict among our motives is an inevitable liability of richness of motive and interests. Striving to be true to the goals one has set for oneself in the past makes possible much of what bears great value in human life, but consistency has a cost. Much of the Daoist text *Zhuangzi* celebrates a life of spontaneity and of concentration on the present moment because a plan-filled, goal-oriented life is one that becomes rigid and creates a filter on one's perceptions of what there is and what is to be valued in the present situation. Those who are highly goal-oriented tend to see their present situation only in terms of what will satisfy their goals, and not in terms of those things in their present situation that might satisfy and stimulate but fail to fit with those goals.[39] One might agree that the nature of human commitments to projects requires a conception of oneself as unified over certain spans of time, but also point out that a project need not extend indefinitely into the future. A young man of twenty-two might commit to trying to make a go of playing in a rock band until he reaches thirty, but leave the future open after that. He thinks of his life as episodic, as Galen Strawson suggests,[40] more along the lines of a series of projects, not necessarily unified by overarching goals or motives. A well-known

[38] Christine M. Korsgaard, "Personal Identity and the Unity of Agency: A Kantian Response to Parfit," *Philosophy and Public Affairs* 18 (1989): 109–15.

[39] In the first chapter of the *Zhuangzi,* for example, Zhuangzi's logician friend Huizi is given seeds that give birth to enormous gourds. When Huizi tries to find a use for the shells of these gourds, he can only think of using them as water dippers or water containers, but they are too big and heavy for such uses. Failing to find a use for them, he smashes them to pieces. Zhuangzi scolds his friend for having underbrush in his head, pointing out that he could have lashed the shells together to make a kind of raft to ride upon the lakes and rivers. See *Chuang Tzu: The Inner Chapters,* trans. A. C. Graham (Indianapolis, IN: Hackett, 2001), chap. 1.

[40] Galen Strawson, "The 'Self'," in *Models of the Self,* ed. Shaun Gallagher and Jonathan Shear (Exeter, UK: Imprint Academic, 1999), 3–24.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

philosopher, not Strawson, once told me that he did not worry about consistency of theme between his many books. He compared himself to an artist who works on a painting, finishes it, and then goes on to the next without thinking of consistency between them.

I am not denying that the requirements of agency might favor a degree of diachronic unity, but they might not reproduce the kind of unity we are accustomed to ascribing to persons over time, nor would they necessarily produce the familiar reasons of prudence that extend into the future. More promising, I believe, is John Locke's suggestion that the concept of personal identity is forensic in nature.[41] That is, it enables the holding of persons as responsible for their past and future actions, making memory one of the crucial psychological threads connecting life slices. Furthermore, a concept of persons enabling ascriptions of responsibility for past and future actions (linking, for example, the one who forms intentions with the one who acts on them) is crucial for the ways in which cooperation is encouraged through punishment of norm violators. Deterrence is possible only with such unity over time. Such enforcement of moral and other cultural norms also seems to require the person as unified to be embodied. It is difficult, at the very least, to imagine a practical way of identifying and punishing noncooperators who fail to be embodied.

The concept of the person persisting over time is, I believe, a prototype concept. As such a concept, it would include the feature of having a human body that persists over time, an embodied consciousness of the environment, and memories, thoughts, intentions, and psychological traits that are continuous and/or connected in a high degree. The explanation for why the person prototype has this content and not some other content is bound up with the cooperative nature of human life and with what it takes to sustain that nature. It is not just that agents are needed who conceive of certain intentions as their own and therefore as needing to be carried out, but also that embodied agents need to be identifiable and held responsible for their cooperativeness or lack of it.[42]

We run into conceptual puzzles when we consider cases, most of them hypothetical or counterfactual, in which some of the prototype characteristics are missing. It then becomes very unclear what the concept of personal identity implies for these cases as to whether the person has

[41] John Locke, "Of Identity and Diversity," in Locke, *An Essay Concerning Human Understanding* (1690), ed. P. H. Nidditch (New York: Oxford University Press, 1979).

[42] Philip Pettit has raised the question of whether my proposal has a Euthyphro problem: Do we hold a person responsible because he's the same person, or do we judge him to be the same person because he is the one in the position to be held responsible? My answer is that the concept of a person persisting over time evolved because the conditions of human cooperative life require relatively long-term persisting agents to hold responsible (the second possibility). But once that concept of a persisting person is in place, with accompanying bodily and psychological criteria for persistence also in place, we hold particular individuals responsible because they are the same persons who did the things for which we need to hold someone responsible.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

persisted or disappeared or become a different person or two different people, and those who react to the cases appear to put forward a variety of intuitions or have none at all (for example, so-called "fission" cases in which a person splits into two who have the same degree of psychological connectedness and continuity with their ancestor; or teletransportation in which one's body on earth is destroyed but one's psychological properties and processes are preserved and associated with a new body on Mars). I think there is often no single right answer as to whether the person has persisted in such cases.[43] Persons, I believe, do not constitute a natural kind. The concept of persons is constructed at least partly in response to practical considerations of the type I have described. To say the concept is constructed, I again want to make clear, is not to say that the very features of the prototype are somehow fabricated. Rather, it is to say that they were *selected as features* of the person prototype out of practical considerations. These practical considerations then motivate a concept of personal identity over time, which is presupposed by the concept of prudence and whatever reasons are associated with it. However, the practical considerations that go into the formation of the relevant prototype or exemplars may not dictate an answer to whether a person has persisted in circumstances that go far beyond the context in which the prototype or exemplars are intelligible. That the structure of the person concept is responsive to social and practical considerations, in other words, might explain why it is very difficult to extend the answer to situations where the social practicality vanishes.[44]

If I am right in my story of how we become selves persisting over time, the reasons we have to provide for our future selves are not based on a necessary and self-evident conception of practical rationality as embodying the fundamental aim of satisfying our desires. Such an aim presupposes a conception of selves persisting over time that is constructed to

[43] Or consider the perplexities that Bernard Williams raises in "The Self and the Future," *Philosophical Review* 79 (1970): 161–80. On the one hand, if I imagine that A's memories are transferred into B's body, and that A's body is tortured, and if I further imagine that it was my memories that got transferred to B's body, I would feel fortunate. This seems to favor the psychological criterion for personal identity. On the other hand, if I am told that I am going to be tortured tomorrow, but that I will not remember anything leading up to the torture, and that, moreover, my impressions of the past will be quite different from the ones I have now, I will still be quite frightened and will not take any comfort in the psychological discontinuity. This seems to favor the bodily criterion of personal identity. Perhaps these conflicting intuitions are explicable if both the bodily criterion and the psychological criterion are present in the prototype without the kind of weighting of these different features that would decide which one is more fundamental to identity.

[44] I agree with David Shoemaker's view, raised in discussion of this paper, that there can be different concepts of personal identity; e.g., biological criteria of identity might be relevant in determining whether I am owed compensation for something that happened to me as a fetus, even if the bodily-psychological concept of the person I have been discussing might not make me the same person as that fetus. Different concepts of the person might arise in different social contexts and on the basis of different practical considerations. And it is possible that different concepts of the person could overlap in their application to the same context and come into conflict. In that case, conceptual revision might be necessary.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

satisfy practical requirements of the social life. These practical requirements are the source of our criteria for judging which desires are desires of our persisting selves and not some other selves, and therefore these requirements are presupposed by the reasons we have for satisfying those desires.

VII. Conclusion

Moral reasons are not given by what will satisfy an individual's desires. Nor are moral reasons simply expressions of social demands on the individual. Rather, they can have a normative grip that reaches deeply into the individual. They do so by referring to situational features that call for actions that support the cooperative nature of human life. Through moral education and socialization, moral reasons become embedded in the individual's motivational propensities and thereby become part of what constitutes that individual as a particular person. The requirements of social cooperation also shape the concept of personal identity over time and the reasons of prudence associated with such a concept. While many have taken reasons of prudence to be the paradigm of reasons instrumentally based on desires of the individual, reasons of prudence, like moral reasons, are rooted in normative demands (based in the cooperative nature of human life) that go into shaping the motivations of the individual.

The story I have told allows that moral reasons can have a deep normative grip on many (perhaps most) individuals, but not necessarily on all. This is the right result, I believe. The normativity of morality is inescapable and can reach deeply into the individual's motivations and not simply answer to whatever the individual's motivations happen to be. However, there will be individuals to whom this inescapability does not matter. These are individuals who are socially constituted, not through moral reasons, but through other kinds of reasons. That is why the common conception is that morality is indeed inescapable, that its normative grip can go deeply into the individual, but also why Thrasymachus is not necessarily irrational for being indifferent to the demands of morality.

The fact that morality has the function of reconciling self-concern and other-concern also explains the conception of moral reasons as overriding reasons. Moral reasons are grounded in what is purported to be appropriate consideration of the agent's self-interest and therefore purport to override reasons of self-interest when such reasons conflict with the appropriate weight given to them in the context of reconciling and balancing self- and other-concern. That is, moral reasons, by virtue of the kind of reasons they are, purport to give due weight to the agent's self-regarding interests and therefore purport to justifiably override those interests when necessary. At the same time, this explanation of the overridingness of moral reasons is consistent with treating the moral perspective as just one

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press

kind of normative perspective[45] that, from the agent's all-things-considered point of view, might not necessarily have overriding force. Moral reasons will fail to have such overriding force if the agent is one in whom such reasons fail to motivationally embed at all, or if the agent cares to some degree about morality but not enough to override other things she cares about.[46] On the one hand, moral reasons purport to override other reasons based on their function of helping to organize the individual's motivational system from the standpoint of promoting and sustaining cooperation. On the other hand, this overriding force is conferred only from that normative perspective. It is a normative perspective that makes human life livable and that goes into the constitution of many or most people, but in the end it can claim superior authority only on the basis of what makes life livable for a community of people. It cannot claim superior authority as a perspective that must motivate each person for that person's life to be livable.

Philosophy, Duke University

[45] Other normative perspectives might include the perspective of self-interest, or of other nonmoral ideals to which the agent has subscribed.

[46] A question from Michael Huemer prompted me to realize the last possibility: that moral reasons might become motivationally embedded but still be overridden by other embedded reasons.

https://doi.org/10.1017/S0265052508080096 Published online by Cambridge University Press