

The Objectivity of Ethics and the Unity of Practical Reason*

Katarzyna de Lazari-Radek and Peter Singer

Evolutionary accounts of the origins of human morality may lead us to doubt the truth of our moral judgments. Sidgwick tried to vindicate ethics from this kind of external attack. However, he ended *The Methods* in despair over another problem—an apparent conflict between rational egoism and universal benevolence, which he called the “dualism of practical reason.” Drawing on Sidgwick, we show that one way of defending objectivity in ethics against Sharon Street’s recent evolutionary critique also puts us in a position to support a bold claim: the dualism of practical reason can be resolved in favor of impartiality.

I. THE PROFOUNDTEST PROBLEM OF ETHICS

One way of attacking the objectivity of ethics is to suggest that an understanding of the origins of our moral judgments casts doubt on their reliability. If, for example, our moral judgments result from our upbringing in a particular culture and others brought up in different cultures have contrary moral judgments, this may be seen as discrediting all such judgments. The appearance of Darwin’s theory of evolution gave rise to a distinctive form of this type of critique, resting on the claim that the judgments we hold have evolved to enhance our prospects of surviving and

* For helpful comments, we would like to thank Derek Parfit, Folke Tersman, Gustaf Arrhenius, and Krister Bykvist, as well as participants in a discussion of the paper at a seminar at the University Center for Human Values, Princeton University, in Fall 2011. We are also most grateful to two reviewers for *Ethics* and four editors of *Ethics*, as well as to Connie Rosati, for the extraordinary amount of work they put into helping us to improve the paper we first submitted—we owe so much to them that it would be tedious to note individually every point at which we have benefited from their suggestions.

reproducing. If different judgments had improved these prospects, we would have had different moral beliefs.

Such arguments against the reliability of our moral beliefs were familiar to the late Victorian moral philosopher Henry Sidgwick. In *The Methods of Ethics*, Sidgwick argued that no account of the origins of our moral judgments could justify us in denying truth and falsity to ethics as a whole. Despite vindicating ethics from this kind of external attack, however, *The Methods* finishes on a famous note of despair. In searching for rational axioms that would give us guidance about what we ought to do, Sidgwick arrived at two that are, at least potentially, in conflict. The axiom of rational egoism says that each of us ought to aim at her or his own good on the whole, and the axiom of benevolence or utilitarianism tells us to aim at the good of all.¹ Sometimes when I aim at my own good I will also maximize the good of all, but often I will have to choose between my own lesser good and the greatest achievable good of others. Sidgwick calls this “the dualism of practical reason” and says that it is “the profoundest problem of ethics.”² His pessimism about ever resolving it is best expressed in the concluding sentence of the first edition of *The Methods*: “But the Cosmos of Duty is thus really reduced to a Chaos: and the prolonged effort of the human intellect to frame a perfect ideal of rational conduct is seen to have been fore-doomed to inevitable failure.”³ Although in later editions he moderated his language, he never changed his mind about the gravity of the problem. The seventh and final edition still concludes with the statement that the dualism of practical reason demonstrates the existence of an “ultimate and fundamental contradiction in our apparent intuitions of what is Reasonable in conduct,” and this contradiction forces us to admit that “the apparently intuitive operation of the Practical Reason, manifested in these contradictory judgments, is after all illusory.” If we are not thereby required to “abandon morality altogether,” we will at least have to abandon “the idea of rationalising it completely.”⁴

In the years since Sidgwick wrote those lines, his “profoundest problem” has neither been resolved nor diminished in significance. Derek Parfit discusses the dualism of practical reason in *On What Matters*. According to what Parfit calls a “wide value-based objective view,” “when one of our two possible acts would make things go in some way that would be impartially better, but the other act would make things go better either for ourselves or for those to whom we have close ties, we often have sufficient reasons to act in either of these ways.”⁵ Parfit’s inclusion of the

1. For discussion of Sidgwick’s acceptance of this axiom, see n. 40, below.

2. Henry Sidgwick, *The Methods of Ethics*, 7th ed. (London: Macmillan, 1907), 386 n. 4. Unless otherwise noted, all further references to *The Methods of Ethics* are to this edition.

3. Henry Sidgwick, *The Methods of Ethics*, 1st ed. (London: Macmillan, 1874), 473.

4. Sidgwick, *Methods of Ethics*, 508.

5. Derek Parfit, *On What Matters*, vols. 1 and 2 (Oxford: Oxford University Press, 2011), 1:137.

word “often” in this sentence marks one difference between his position and Sidgwick’s. Parfit thinks that in extreme cases—for example, when I can save a stranger’s life at a trivial cost to myself—it would be irrational to act solely on the basis of what is better for myself. Nevertheless, he agrees with Sidgwick that when doing what is right would require a serious sacrifice of one’s own interests, it is not irrational to do what is in one’s own interests. In contrast to Sidgwick, Parfit does not see the problem as an indication of the collapse of practical reason, but he does regard it as showing the limits of practical reason: in many cases reason is not a perfect guide because it does not lead us in a single direction.

What might these two problems—the evolutionary critique of objectivity in ethics and the dualism of practical reason—have in common? Surprisingly, Sidgwick’s response to the evolutionary critique of objectivity in ethics points us toward a way of overcoming the dualism of practical reason. We are not the first to notice that Sidgwick provided a compelling answer to claims that a sound understanding of the origins of our moral beliefs gives us reasons for abandoning objectivity in ethics.⁶ While Sidgwick successfully responded to such evolution-based arguments against the objectivity of ethics, however, he curiously failed to see—as have commentators on *The Methods of Ethics*—that his response provided a basis for a solution to his own, very different, worry about the objectivity of ethics.

We begin this article by drawing on Sidgwick’s own defense of objectivity in the face of theories about the origins of our moral judgments. We then move to Sharon Street’s forceful restatement of the view that an evolutionary understanding of human nature is incompatible with moral realism.⁷ We argue that Sidgwick’s defense of objectivity in ethics survives Street’s restatement of the argument, but we will then add that this same understanding of evolution suggests that practical reason, freed from some specific distortions that derive from our evolutionary origins, ceases to be divided between partial and impartial modes of reasoning. Thus, we make use of Sidgwick’s own argument to defend a claim that goes beyond anything Sidgwick himself suggested but which we believe he would have welcomed: the dualism of practical reason can be resolved in favor of impartiality. Our case for this claim has to take this indirect route because our defense of objectivity in ethics against an evolutionary critique is crucial to our resolution of the dualism. Only when objectivism has been defended in a particular way do we find ourselves with a general view of truth in ethics from which the resolution of the dualism follows as a specific application.

6. Hallvard Lillehammer, “Methods of Ethics and the Descent of Man: Darwin and Sidgwick on Ethics and Evolution,” *Biology and Philosophy* 25 (2010): 361–78.

7. Sharon Street, “A Darwinian Dilemma for Realist Theories of Value,” *Philosophical Studies* 127 (2006): 109–66.

II. SIDGWICK AND THE ARGUMENT FROM THE ORIGINS OF OUR MORAL FACULTY

In defending the project of *The Methods of Ethics*, Sidgwick considers the view that knowledge of the origins of our moral intuitions may show that they cannot be valid, that is, that they are not a reliable indication of the truth of what we intuit.⁸ He rejects this view, arguing that if an intuition seems to be self-evident, we should not regard it as untrustworthy merely because it was “caused in known and determinate ways.” He goes further still, denying that those who affirm the truth of judgments that we intuitively grasp as self-evident need even demonstrate that the causes of these judgments are of a kind that is likely to lead to true judgments. That requirement would, Sidgwick argues, lead to a kind of infinite regress that would make it impossible ever to find certainty about anything, for “the premises of the required demonstration must consist of caused beliefs, which as having been caused will equally stand in need of being proved true, and so on *ad infinitum*.”⁹ In other words, if *all* our beliefs are equally the effect of some prior causes, this fact alone cannot give us grounds to reject them, for these grounds would apply with equal force to all of our beliefs, thus committing us to total skepticism about everything. Moreover, Sidgwick points out, no theory of the origins of our moral judgments can show that the fundamental ethical conceptions of “right” or “what ought to be done” or “good” or “what it is reasonable to desire or seek” are invalid and that therefore *all* propositions of the form “X is right” or “X is good” are untrustworthy because these propositions are about something fundamentally distinct from the subject matter of the sciences.

Although Sidgwick firmly defends ethics from a general skeptical attack grounded on any theory of the origins of our moral intuitions, he acknowledges that a more limited claim could be successful: “It may, however, be possible to prove that some ethical beliefs have been caused in such a way as to make it probable that they are wholly or partially erroneous.” He adds that it will “hereafter” be important to consider whether any ethical intuitions that we are disposed to accept as valid are open to attack on such grounds but that his “present” concern is only to deny the more general argument against the trustworthiness of the moral faculty.¹⁰ We will follow Sidgwick in separating the general objection to ethics as a

8. Sidgwick makes no explicit reference to Darwin in any edition of *The Methods*. That he has Darwin’s account of the origins of our moral intuitions in mind, however, is confirmed by the fact that he makes a very similar point in a separate essay titled “The Theory of Evolution in Its Application to Practice,” published in the first volume of *Mind*, in 1876. (We owe this reference to Lillehammer, “Methods of Ethics and the Descent of Man.”)

9. Sidgwick, *Methods of Ethics*, 213.

10. *Ibid.*

whole, based on theories about the origins of our moral judgments, from more specific objections to particular ethical beliefs, and we will discuss them in turn.¹¹

III. THE GENERAL OBJECTION AND STREET'S DARWINIAN DILEMMA

Since Sidgwick's day, and especially over the past forty years, an extensive literature has developed on the origins of morality and of our moral intuitions, much of it informed by a considerable body of empirical research.¹² It is not surprising that this body of theory should lead to further discussion of the implications of our understanding of evolutionary theory for morality. The most widely discussed and philosophically sophisticated contemporary argument for the kind of view that Sidgwick rejected is Sharon Street's claim that a "Darwinian Dilemma" faces those who hold a realist theory of value.¹³ Street starts from a premise that we fully accept: "Evolutionary forces have played a tremendous role in shaping the content of human evaluative attitudes."¹⁴ She then argues that those who defend objective moral truth face a choice between two uncongenial possibilities. The first possibility is that evolutionary forces have no tendency to lead to the selection of beings who hold objectively true evaluative attitudes. In this case, objectivists will have to admit that most of our evaluative judgments are unjustified. The second possibility is that evolutionary forces did favor the selection of those who are able to grasp objective moral truths. But this, Street argues, is contrary to a scientific understanding of how evolution works.

To take the first horn of the dilemma and accept that evolutionary forces have no relation to objectively true evaluative attitudes means,

11. Many subsequent nonnaturalists gave arguments similar to Sidgwick's for the limited relevance to ethics of knowledge of the origin of our moral ideas. See, e.g., Hastings Rashdall, *The Theory of Good and Evil* (Oxford: Clarendon, 1907), vol. 2, chap. 4, esp. secs. 1 and 9; W. D. Ross, *The Foundations of Ethics* (Oxford: Clarendon, 1939), 15–17; and C. D. Broad, "Critical Notice of Julian Huxley's *Evolutionary Ethics*," *Mind* 53 (1944): 344–67, repr. in *Broad's Critical Essays in Moral Philosophy*, ed. David Cheney (London: Allen & Unwin, 1971), 156–87, esp. 169–72, 178–80.

12. See, e.g., W. D. Hamilton, "The Genetical Evolution of Social Behaviour," pts. 1 and 2, *Journal of Theoretical Biology* 7 (1964): 1–16, 17–52; R. L. Trivers, "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology* 46 (1971): 35–57; James Q. Wilson, *The Moral Sense* (New York: Free Press, 1993); Robert Wright, *The Moral Animal* (New York: Pantheon, 1994); Matt Ridley, *The Origins of Virtue* (London: Viking, 1996); and Jonathan Haidt, *The Righteous Mind* (New York: Pantheon, 2012). Also relevant is J. D. Greene, "The Secret Joke of Kant's Soul," in *Moral Psychology*, vol. 3, *The Neuroscience of Morality: Emotion, Disease, and Development*, ed. W. Sinnott-Armstrong (Cambridge, MA: MIT Press, 2007), 35–80.

13. Street, "Darwinian Dilemma," 109–66.

14. *Ibid.*, 109, 122.

Street suggests, that our prospects of having evaluative attitudes that lead us to moral truths are like the prospects of sailing to Bermuda while allowing our boat's course to be determined by the winds and tides. We would be incredibly lucky to reach Bermuda, and if we did, it would be a remarkable coincidence. Barring such a coincidence, however, the realist has to accept what Street considers a "far-fetched skeptical result," namely, that "most of our evaluative judgments are off-track due to the distorting pressure of Darwinian forces."¹⁵

Those taking the second horn of the dilemma fare no better. They make a claim that is unacceptable on scientific grounds. Street offers a list of some of the judgments we make, which includes, for example, "We have greater obligations to help our own children than we do to help complete strangers."¹⁶ Such judgments are conducive to reproductive success, so it is easy to see how evolutionary forces would lead us to make them. It is not so easy to see how evolutionary forces would lead us to make only judgments that are objectively true. Why should the truth of a judgment be something that evolution favors? As Street says, it is more scientifically plausible to explain human evaluative attitudes as having evolved because they help us to survive and to have surviving offspring than because they are true.

To show how evolution could shape our evaluative judgments, Street asks us to suppose that we had evolved as a different kind of being. Social insects, for example, have a stronger orientation toward the welfare of the community than to their own individual survival, and male lions kill offspring that are not their own. Assuming that in some way we could be intelligent, but with reproductive patterns more like those of social insects or lions, we would, she claims, have different basic evaluative attitudes that would lead us to make different reflective evaluative judgments. Since not all these judgments could be true, wouldn't it be a remarkable coincidence if we just happened to have evolved as the kind of beings that make true evaluative judgments?¹⁷

Street's speculation about intelligent social insects echoes one that Darwin made in *The Descent of Man*, when he wrote that if we were reared in conditions like beehives, our unmarried females would "think it a sacred duty to kill their brothers, and mothers would strive to kill their fertile daughters; and no one would think of interfering."¹⁸ When Frances Power Cobbe, in *Darwinism in Morals, and Other Essays*, lamented that this

15. *Ibid.*, 109.

16. *Ibid.*, 115.

17. *Ibid.*, 120.

18. Charles Darwin, *The Descent of Man* (London: Penguin, 2004), 122. (We owe this reference to Lillehammer, "Methods of Ethics and the Descent of Man.")

view “aims . . . a deadly blow at ethics,” Sidgwick responded that this was not so. Cobbe had overlooked, Sidgwick wrote, the ability of the principle of utilitarianism to accept “almost any degree of variation in actual rules” without giving up the idea of absolute moral duties.¹⁹ Thus, Sidgwick appears to accept that our particular set of moral rules does not hold for all ways of living, but he denies that the truth of Darwin’s claim could undermine our belief that there are some things that we ought to do. In very different circumstances, killing one’s brother or daughter could be the right thing to do. What would remain constant, however, is an ultimate principle like “Do what is best for the well-being of all.” Hence ethics as a whole is not threatened by the theory of evolution.

This suggests that Sidgwick, if armed with a modern understanding of evolutionary psychology, could reach a verdict not far from what Street describes as the “far-fetched skeptical result” that “most of our evaluative judgments are off-track due to the distorting pressure of Darwinian forces.” Sidgwick and Street could argue about whether it is “most” or only “many” of our common moral judgments that are offtrack, but a contemporary Sidgwick might be closer to Street than the historical one. Sidgwick thought that what he calls “the morality of common sense”—that is, the set of moral rules that we intuitively assume to be true—tends to produce actions that maximize utility. But if he shared our modern scientific understanding that evolutionary forces operate at the level of the gene or the individual, or at most the community, rather than at the level of the species (and certainly not at the level of all sentient beings), he would surely have been open to the possibility that these evolutionary forces have produced evaluative attitudes that fail to conduce to ultimate moral truths such as “Do what is best for the well-being of all.” To the extent that our common moral judgments are affected by these evolutionary forces, it would then have been consistent with Sidgwick’s own approach to the morality of common sense for him to reject the particular judgments to which these forces led, while maintaining the validity of the more general principle that we should do what is best for the well-being of all. This is, after all, what he already does in Book III of *The Methods* with many of the particular judgments of commonsense morality.²⁰

19. H. Sidgwick, review of *Darwinism in Morals*, by Frances Power Cobbe, *Academy*, June 15, 1872. (Once again, we owe this reference to Lillehammer, “Methods of Ethics and the Descent of Man.”)

20. In saying that Sidgwick would have rejected these judgments, we mean that he would not have taken them to state true moral principles. Whether they should continue to be included among the set of moral rules that people are encouraged to follow would, for Sidgwick, depend on whether continuing to include them would have better consequences than dropping them.

The position we have just attributed to Sidgwick avoids Street's dilemma by accepting its first horn, for many of our common moral judgments. Street would no doubt then try to press her argument against the ultimate principle. How do we reach it, if it has no relation to our evolved basic evaluative attitudes? Was it sheer coincidence, like our drifting boat reaching Bermuda? When it comes to an ultimate principle like doing what is best for the well-being of all, however, Sidgwick has a good response to this argument. He believes we come to understand such principles by the use of our reason.²¹

At this point Sidgwick could take the second horn of Street's dilemma. Street focuses on the question of whether evolution is likely to lead us to have a capacity to recognize objective moral truths. If our moral beliefs are evolutionarily advantageous, then the advantages they confer on us in surviving and reproducing have nothing to do with their truth. So why would evolution have led us to have a capacity to recognize moral truth? Street correctly points out that a specific capacity for recognizing moral truths would not increase our reproductive success. But a capacity to reason *would* tend to increase our reproductive success. It may be that having a capacity to reason involves more than an ability to make valid inferences from premises to conclusions. It may include the ability to recognize and reject capricious or arbitrary grounds for drawing distinctions and to understand self-evident moral truths—what Sidgwick referred to as “rational intuition.” In other words, we might have become reasoning beings because that enabled us to solve a variety of problems that would otherwise have hampered our survival, but once we are capable of reasoning, we may be unable to avoid recognizing and discovering some truths that do not aid our survival.²² That can be said about some complicated truths of mathematics or physics. It can also, Parfit has suggested, be the case with some of our normative epistemic beliefs, for instance, the belief that,

21. Parfit, *On What Matters*, esp. Pts. One and Six, offers a forceful and detailed defense of a position very like Sidgwick's.

22. Colin McGinn suggests this explanation of why evolution has not eliminated moral behavior in “Evolution, Animals and the Basis of Morality,” *Inquiry* 22 (1979): 91. One of us has defended a similar view in Peter Singer, *The Expanding Circle* (Princeton, NJ: Princeton University Press, 2011; first published 1981), chap. 5, and also in “Ethics and Intuitions,” *Journal of Ethics* 9 (2005): 331–52. (These writings do not defend objectivity in ethics, as we do here, nor do they indicate that this form of argument could be significant for overcoming the dualism of practical reason.) Parfit makes a related argument against Street in *On What Matters*, 2:492–97. A broader application of this view can be found in Steven Pinker, *The Better Angels of Our Nature* (New York: Viking, 2011), where Pinker draws on *The Expanding Circle* as well as history, psychology, and cognitive science to make the case that our capacity to reason (which he sees as having been enhanced by the invention of printing and subsequent social developments) is partly responsible for the decline in violence in recent times. See esp. Pinker, *Better Angels*, 642–70, 689–92.

when some argument is valid and has true premises so that this argument's conclusion must be true, these facts give us a decisive reason to believe this conclusion. Parfit argues that this normative claim, about what we have decisive reason to believe, is not itself evolutionarily advantageous since to gain that advantage, it would have been sufficient to have the nonnormative beliefs that the argument is valid, and has true premises, and that the conclusion must be true. Hence, this and other normative epistemic beliefs are not open to a debunking argument.²³ This may also hold for some of our moral beliefs. One such moral truth could be Sidgwick's axiom of rational benevolence: "each one is morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him."²⁴

It may be objected that if some aspects of our capacity to reason conferred an evolutionary advantage, while other aspects were disadvantageous in that respect (perhaps because they lead us to act more altruistically than we would otherwise have done), then these other aspects would have been selected against and would have disappeared. (They might also have disappeared even if they were merely neutral, neither advantageous nor disadvantageous, because of evolutionary drift, but obviously the more a trait or capacity disadvantages the being who possesses it, the more rapidly it is likely to disappear.) It appears to be the case, however, that we have retained capacities to reason that do not confer any evolutionary advantage and may even be disadvantageous. How can that be? A plausible explanation of the existence of these capacities is that the ability to reason comes as a package that could not be economically divided by evolutionary pressures. Either we have a capacity to reason that includes the capacity to do advanced physics and mathematics and to grasp objective moral truths, or we have a much more limited capacity to reason that lacks not only these abilities but others that confer an overriding evolutionary advantage. If reason is a unity of this kind, having the package would have been more conducive to survival and reproduction than not having it.

Street discusses the objection that our capacity to grasp objective moral truths could be a by-product of some other evolved capacity. She argues that this capacity must be a highly specialized one "specifically attuned to the evaluative truths in question."²⁵ Therefore, those who make

23. Parfit, *On What Matters*, 2:492, and e-mail to the authors, August 16, 2011.

24. Sidgwick, *Methods of Ethics*, 382. Although Sidgwick uses the term "rational benevolence" to describe his axiom, we will also refer to it as the principle of "universal benevolence." We will argue that we reach this principle through reason, but we prefer to avoid terminology that prejudges that question.

25. Street, "Darwinian Dilemma," 143.

this proposal face the Darwinian dilemma once again, this time with respect to the relationship between the specialized capacity to grasp objective moral truths and the other more basic evolved capacity. Either there is no relationship between the evolution of the basic capacity and the independent moral truths—in which case it is a remarkable coincidence that the basic capacity had, as a by-product, a capacity to grasp objective moral truths—or there is some relationship between the evolved basic capacity and the capacity to grasp independent moral truths. We have taken the second horn of this dilemma. Those who take this course, Street says, must claim that the evolved capacity “involves at least some *basic* sort of ability to grasp independent evaluative truths, of which our present-day ability to grasp evaluative truths is a refined extension, in much the same way that our present-day ability to do astrophysics is presumably a refined extension of more basic abilities to discover and model the physical features of the world around us.” She then adds, “But at this point the realist has to give some account of how this more basic sort of ability to grasp independent evaluative truths arose.”²⁶ Indeed, that is true, but given that philosophers like Sidgwick have long said that it is our capacity to reason that enables us to grasp moral truths, and given that we can explain why a capacity to reason would have been evolutionarily advantageous, it is odd that Street does not directly confront the idea that the capacity to grasp moral truths is simply an application of our capacity to reason, which enables us to grasp *a priori* truths in general, including both the truths of mathematics and moral truths. For if the ability to grasp moral truths is an aspect of our ability to reason, and to respond to reasons, it is easy to give an account of how it arose.

IV. THE PARTICULAR OBJECTION: HOW UNIVERSAL BENEVOLENCE SURVIVES THE EVOLUTIONARY CRITIQUE

That is, however, not all that Sidgwick might say in reply to Street. We noted earlier that he postponed discussion of the possibility that some particular ethical beliefs have been caused in such a way as to make it probable that they are mistaken. Since we can now construe Street’s argument as limited to an attack on a particular belief, specifically that we ought to maximize well-being generally, we should note what he says about such an argument. The discussion of this point comes toward the end of Book III of *The Methods*, when Sidgwick has presented the three moral principles or axioms, the principles of justice, prudence, and benevolence, that he takes to be self-evident. After thus satisfying himself and (he hopes) the reader that we can know some moral truths by intuition, he explains why

26. *Ibid.*, 144.

he earlier refrained from a lengthy discussion of the origins of our moral intuitions. The reason is, he says, that no theory of the origins of our moral intuitions “has ever been put forward professing to discredit the propositions that I regard as really axiomatic, by showing that the causes which produced them were such as had a tendency to make them false.” On the other hand, an argument that targeted any of our other moral intuitions on these grounds in order to show that they are not absolutely true would be superfluous since the kind of direct reflection that has occupied him for most of Book III has already led to this conclusion. Finally, Sidgwick adds, if a theory of the origins of our moral rules viewed them as existing because they are broadly means to the ends of improving the welfare either of individuals or of the larger community, then this would tend to confirm the results that he has reached by a different method since they show that the rules of commonsense morality are subordinate to the principles of prudence and benevolence.²⁷

We can now ask: Is it still true, after all the work that has been done on the origins of our moral intuitions since Sidgwick’s time, that no theory has been put forward professing to discredit the propositions that he regards as really axiomatic? Remarkably, we believe that it is, at least for the all-important axiom of universal benevolence. After all, that axiom contradicts the very evaluative attitudes that Street offers as examples of judgments that are likely to lead to reproductive success, such as “We have greater obligations to help our own children than we do to help complete strangers.” Evolutionary theorists have long had difficulty in explaining how pure altruism is possible. They tend to explain it in terms of more limited forms of altruism, such as altruism toward kin and reciprocal altruism, that is, altruism toward those with whom we are in a cooperative relationship. Some theorists also accept the possibility of altruism toward one’s own group. It is, however, difficult to see any evolutionary forces that could have favored universal altruism of the sort that is required by the axiom of rational benevolence. On the contrary, there are strong evolutionary forces that would tend to eliminate it.

There is a popular misconception that altruism can arise because it is “for the good of the species.” Modern evolutionary theorists point out that while species go in and out of existence only over very long periods of time, individuals are much more short-lived. This means that individuals who behave altruistically would be likely to be selected against, and eliminated from the population, before they could become common enough to have any impact on the survival of the species as a whole. Richard Dawkins has argued—as the title of his early work, *The Selfish Gene*, suggests—that actions that involve sacrificing an organism’s prospects of surviving and reproducing have evolved because they benefit the organ-

27. Sidgwick, *Methods of Ethics*, 383–84.

ism's genes, largely through favoring kin. He does not hesitate to draw the conclusion that "much as we might wish to believe otherwise, universal love and the welfare of the species as a whole are concepts that simply do not make evolutionary sense."²⁸ Pierre van den Berghe has said flatly, and no doubt too bluntly, that "we are programmed to care only about ourselves and our relatives."²⁹ Richard Alexander, in *The Biology of Moral Systems*, writes: "I suspect that nearly all humans believe it is a normal part of the functioning of every human individual now and then to assist someone else in the realization of that person's own interests to the actual net expense of those of the altruist. What this greatest intellectual revolution of the century [i.e., the individualistic perspective in evolutionary biology] tells us is that, despite our intuitions, there is not a shred of evidence to support this view of beneficence, and a great deal of convincing theory suggests that any such view will eventually be judged false."³⁰

In *Unto Others*, Elliot Sober and David Sloan Wilson have forcefully challenged this individualistic perspective in evolutionary theory. They argue that evolution could have selected for actions that benefit groups to which individuals belong, rather than for actions that benefit the individuals themselves. For the argument we are about to make, therefore, it is vital to understand that, while Sober and Wilson are challenging the views of Dawkins, van den Berghe, and Alexander, they do *not* argue that evolution could have selected for the kind of universal benevolence required by Sidgwick's axiom. As they put it, "our goal in this book is not to paint a rosy picture of universal benevolence. Group selection does provide a setting in which helping behavior directed at members of one's own group can evolve; however it equally provides a context in which hurting individuals in other groups can be selectively advantageous. Group selection favors within-group niceness *and* between-group nastiness."³¹ In the absence of an appeal to our evolved capacity to reason as the basis for our ability to grasp moral truth, therefore, it is difficult to see what plausible

28. Richard Dawkins, *The Selfish Gene* (Oxford: Oxford University Press, 1976), 2. For a summary of the debate on this issue, see Samir Okasha, "Biological Altruism," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (Stanford, CA: Stanford University, 2009), <http://plato.stanford.edu/archives/win2009/entries/altruism-biological>.

29. Pierre L. van den Berghe, "Bridging the Paradigms: Biology and the Social Sciences," in *Sociobiology and Human Nature*, ed. M. S. Gregory, A. Silvers, and D. Sutch (San Francisco: Jossey-Bass, 1978), 32–52.

30. Richard Alexander, *The Biology of Moral Systems* (New York: de Gruyter, 1987), as quoted by Elliot Sober and David Sloan Wilson, *Unto Others* (Cambridge, MA: Harvard University Press, 1998), 5–6. We are not endorsing Alexander's account of human nature but merely indicating that if there is such a phenomenon as universal benevolence, it is not easy to explain—without appealing to the rationality of the principle of universal benevolence—how the phenomenon evolved.

31. Sober and Wilson, *Unto Others*, 9.

evolutionary explanation there could be for the idea of equal concern for the interests of complete strangers who do not belong to one's own group. Moreover, even if "altruism for the good of the species" somehow were the product of our evolution, even that would not go far enough, for the principle of universal benevolence bids us to have concern not only for the good of our own species but for all sentient beings.³²

Street argues that an evolutionary account of the origins of morality is incompatible with moral realism. We have seen that Sidgwick has good reasons for denying that such an argument undermines the normative truth that each of us ought to give as much weight to the good of anyone else as we give to our own good. On the other hand, an evolutionary understanding of the origins of our ethical judgments does seem to undermine some of our ethical judgments, at least to the extent of suggesting that we should not take them for granted merely because we intuitively judge them to be sound.³³

Consider, for instance, the judgment that incest is wrong, even when those involved are adult siblings. Among our ancestors, for millions of years, such sexual relationships probably increased the proportion of abnormal offspring and, hence, diminished prospects of reproductive success, as compared to sexual relationships between those who were not so closely related. Hence, our negative evaluative attitude toward incest—which is less universally held when the degree of consanguinity, and hence the risk of abnormal offspring, is reduced—is easily explained as part of our evolutionary heritage. But today it is possible to separate sex and reproduction, so this reason for rejecting incest in the circumstances described is no longer always applicable. Thus the judgment that incest is always wrong can be seen to be the product of a cause that, in at least some cases, produces judgments likely to be in error.³⁴ Something similar may be true of the widespread, although not universal, attitude that homosexuality is wrong since it is even less likely than incest to lead to reproductive success.

Roger Crisp, in *Reasons and the Good*, offers a further example of an intuition for which an evolutionary explanation is available: "On Monday I blind a stranger to prevent his buying the last copy of a CD I want to buy.

32. Sidgwick, *Methods of Ethics*, 414.

33. Guy Kahane puts it like this: "if only some of our evaluative beliefs are susceptible to the relevant kind of evolutionary explanation, and we can at least roughly gauge the degree of this evolutionary influence on various beliefs, then what we should get isn't evaluative skepticism but a proportional lowering of justification" ("Evolutionary Debunking Arguments," *Noûs* 45 [2011]: 103–25, quote at 119).

34. See Jonathan Haidt, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108 (2001): 814–34, citing Jonathan Haidt, Fredrik Björklund, and Scott Murphy, "Moral Dumbfounding: When Intuition Finds No Reason" (unpublished manuscript).

I buy the CD. On Tuesday I buy another CD, knowing that I could have given the money to Sight Savers International and prevented the blindness of at least one person.” Crisp points out that most people would think that the blinding is wrong, but the failure to prevent blinding is not, although the consequences of both are largely the same. He adds that this is the kind of morality that we would expect to result from evolution because “it is clear that a group cannot function well if its members are permitted to harm one another, whereas the survival value of a prohibition on allowing others to suffer is more dubious.”³⁵

As we have already mentioned, Street uses as an example of how our judgments coincide with intuitions likely to lead to reproductive success, the judgment “We have greater obligations to help our own children than we do to help complete strangers.” The common intuition that this judgment is true may be the result of the fact that those who accept it would be more likely than those who do not accept it to leave surviving offspring to carry on their genes. In reviewing commonsense morality, Sidgwick writes that when we consider the duty of parents to their children as such, without taking into consideration psychological and social aspects of how best to bring up children, it is not at all self-evident “that we owe more to our own children than to others whose happiness equally depends on our exertions.” In support of this view, Sidgwick asks us to imagine that my family and I land on a desert island where I find an abandoned child. It is not, he seems to think, self-evident that I have a lesser obligation to provide for the subsistence of this child than I do to provide the same for my own children.³⁶ This is not to say that the judgment that we have greater obligations to help our own children than to help strangers cannot be justified but rather that if it is to be justified, it needs a form of justification that does not start from the idea that because we strongly feel that it is right, it must be true. For instance, it may be the case that our nature is such that the most reliable way of raising happy, well-adjusted children is to raise them in a close, caring family united by natural ties of love and affection. If so, then this would provide an indirect justification of the judgment that we have greater obligations to our own children than to the children of strangers. Given the kind of creatures we are—not social insects but mammals with children who are dependent on us for many years—loving our own children and helping them more than we help the

35. Roger Crisp, *Reasons and the Good* (Oxford: Clarendon, 2006), 21. Gilbert Harman offers a similar explanation of the distinction—with the more specific proviso that moral attitudes derive from implicit agreement, and whereas everyone would benefit from an agreement not to harm others, the rich and strong would not benefit from an agreement to help others. See his “Moral Relativism Defended,” *Philosophical Review* 84 (1975): 3–22, quote at 12.

36. Sidgwick, *Methods of Ethics*, 346–47.

children of strangers would, on this view, be justified in terms of the more ultimate principle that Sidgwick mentioned, that it is good to do what is best for the well-being of all.

Crisp accepts what he calls the “Self-Interest Principle,” which states, in effect, that any agent has a reason to do what makes her life go better, the strength of the reason varying in proportion to the extent to which her well-being will be improved.³⁷ But this principle can itself be debunked, in much the way that Crisp debunked the acts and omissions distinction. As Folke Tersman writes (using “SI” to stand for the Self-Interest Principle):

A debunking explanation of SI can be given along the following lines. It is safe to assume that at least some concern for one’s self-interest is the result of evolutionary pressure, and the conviction that we have a reason to act self-interestedly can be seen as a way of verbalizing that concern, given the role of such judgments in planning and deliberation. The universal element of SI—the part that entails that it holds for everyone—needs another explanation. But then we can appeal to the cognitive processes mentioned above. We search for generality and coherence, and try to find relevant similarities and ignore irrelevant differences. If we restrict the scope of SI, we need an explanation in terms of relevant differences between the persons for whom it holds and those for whom it does not hold. The universal version does not require such complexity, and is therefore attractive for the reflective mind that seeks simplicity. So, the fact that reflection on SI can prompt us to accept it comes as no surprise.³⁸

Tersman’s point is that the fact that a cognitive process is involved in the formation of an intuition does not show that the intuition cannot be debunked. Just as we cannot trust the conclusion of a valid deductive argument if it starts from premises not known to be true, so we cannot trust the conclusion of an intuition reached by a cognitive process unless we know that the starting point of the process is true. Guy Kahane makes a similar point against the claim that one of us (Peter Singer) has previously made that an evolutionary debunking argument strengthens the case for utilitarianism. Kahane says that if evolution has selected for a disposition to altruism toward one’s kin and those with whom one is in reciprocal relationships, then we should suspect not only principles that support altruism toward kin and cooperating partners but also the “reasoned extension of such partial forms of altruism to universal altruism.” Otherwise, he says, we risk being like a person who comes upon a madman counting the blades of grass in his backyard and tells him that

37. Crisp, *Reasons and the Good*, 73.

38. Folke Tersman, “The Reliability of Moral Intuitions: A Challenge from Neuroscience,” *Australasian Journal of Philosophy* 86 (2008): 389–405, 403.

because the distinction between his backyard and anyone else's backyard is arbitrary, he should instead be counting blades of grass everywhere in the world.³⁹

We accept that if a starting point can be debunked, it cannot lend support to a more general or less arbitrary version of itself. But in *The Methods* Sidgwick did not develop the case for his axiom of universal benevolence by arguing for a reasoned extension of egoism or partial altruism. Instead, he claimed that it is self-evident that a mere difference in time does not give some moments of our own existence greater significance than any other moments. This is, for Sidgwick, the "self-evident element" in the principle of prudence, and he goes on to argue that the reasoning which enables us to see this as self-evident also enables us to see as self-evident that "the good of any one individual is of no more importance, from the point of view (if I may say so) of the Universe, than the good of any other." He then adds that it also seems to him to be self-evident that "as a rational being I am bound to aim at good generally,—so far as it is attainable by my efforts,—not merely at a particular part of it."⁴⁰

Thus, we can agree with Tersman's debunking explanation of the Self-Interest Principle. Crisp's Self-Interest Principle extends the idea that

39. Kahane, "Evolutionary Debunking Arguments," 119, referring to Singer, "Ethics and Intuitions," 350–51.

40. Sidgwick, *Methods of Ethics*, 381–82. There has been an extensive discussion over whether what Sidgwick calls the "axiom of prudence" expresses the idea of egoism. The form of the axiom that he offers in Bk. III, Chap. xiii, is, as we noted above, about having equal concern for all moments of our existence over time and not about preferring our own good to the good of others (381, 383). In the autobiographical preface to the sixth edition, however, Sidgwick has in mind a different principle of egoism when he says that "the rationality of self-regard seemed to me as undeniable as the rationality of self-sacrifice." It is to this form of rational egoism that he refers when, in the concluding chapter he sets up the problem of the dualism of practical reason (xviii, 497–98). Parfit has suggested (e-mail to the authors, June 25, 2012) that the fact that Sidgwick does not include the principle of rational egoism among the axioms he discusses in Bk. III, Chap. xiii, could be taken as evidence that he grasped that rational egoism is in some way less secure or undeniable than the axioms he does endorse in that chapter. This leads us to a view slightly different from that taken by J. B. Schneewind, who argues that to understand why Sidgwick believed that there is a dualism of practical reason, we should understand the axiom of prudence as "my own greatest happiness is the rational ultimate end for me" (*Sidgwick's Ethics and Victorian Moral Philosophy* [Oxford: Clarendon, 1977], 290–97). We see this as a distinct principle and not the "axiom of prudence" defended in Bk. III, Chap. xiii. Doubts about whether Sidgwick's axiom of prudence is an axiom of egoism go back as far as G. von Gizycki, review of *The Methods of Ethics*, by Henry Sidgwick, *Ethics* 1 (1890): 120–21, and more recently have been expressed by R. Shaver, *Rational Egoism* (Cambridge: Cambridge University Press, 1999), 74–77; B. Schultz, *Eye of the Universe* (Cambridge: Cambridge University Press, 2004), 213; and J. Skorupski, "Three Methods and a Dualism," in *Henry Sidgwick*, ed. R. Harrison (New York: Oxford University Press, 2001), 61–82. See also David Phillips, *Sidgwickian Ethics* (New York: Oxford University Press, 2011), 138–39.

I have a reason to act in my own interests only by granting that others have similar reasons to act in their own interests. This modest extension can be seen as inherent in the very concept of what it is to have a reason—it is the “first stage” of universalization that J. L. Mackie described as “the irrelevance of numerical differences.”⁴¹ Mackie was prepared to accept that this minimal stage of universalization is inherent in the meaning of “ought” and other moral terms. On the other hand, he rejected R. M. Hare’s contention that this notion of universalizability is sufficient to get us to a form of utilitarianism. That, Mackie said, involves a substantive moral claim, not to be found in the meanings of the moral terms or in the bare concept of what it is to have a reason. We agree, and the same is true of the principle of universal benevolence. It rests on a substantive claim, and we cannot get to it from any form of egoism merely by seeking, in Tersman’s words, “generality and coherence” or Kahane’s “reasoned extension.” There is nothing *incoherent* in accepting the principle of self-interest while rejecting the principle of universal benevolence.⁴² Even if there were such an incoherence, however, the fact that, as we have shown, there is another way of reaching the principle of universal benevolence would suffice to establish that it is not founded on a contaminated starting point.

Tersman contends that, to avoid general skepticism about ethics, “one must show that there are intuitions for which no debunking explanation can be given or where the debunking explanations are inferior to non-debunking ones.” He then adds: “Let us say that if an explanation of an intuition entails that it is true or likely then it is ‘validating’. In my view, if an explanation appeals to the way the intuition was formed, it is validating only if combined with an account of *why* the fact that it was so formed makes it true or significantly likely. And that account must both be described in some detail and have some degree of independent plausibility—not just any *ad hoc* story would do.”⁴³ We suggest that this is indeed the case with the principle of universal benevolence. We form the intuition as a result of a process of careful reflection that leads us to take, as Sidgwick puts it, “the point of view of the universe.” This idea is not specific to any particular cultural or religious tradition. On the contrary, the lead-

41. J. L. Mackie, *Ethics: Inventing Right and Wrong* (London: Penguin, 1977), chap. 4.

42. Sidgwick makes a similar point when he rejects John Stuart Mill’s “proof” of utilitarianism, which is based on a generalization from the desire for one’s own happiness. Sidgwick responds: “There being . . . no actual desire, so far as this reasoning goes, for the general happiness, the proposition that the general happiness is desirable cannot be in this way established; so that there is a gap in the expressed argument, which can, I think, only be filled by some such proposition as that which I have above tried to exhibit as the intuition of Rational Benevolence” (*Methods of Ethics*, 388). At 497–98, Sidgwick clearly accepts that egoism is not incoherent.

43. *Ibid.*, 403–4.

ing thinkers of distinct traditions have independently reached a similar principle and have regarded it as the essence of morality. In addition to the well-known Jewish and Christian versions of the Golden Rule, we find similar ideas in the Confucian, Hindu, and Buddhist traditions.⁴⁴ Finally, there is no plausible explanation of this principle as the direct outcome of an evolutionary process, nor is there any other obvious non-truth-tracking explanation. Like our ability to do higher mathematics, it can most plausibly be explained as the outcome of our capacity to reason. Admittedly, the absence of good rival explanations for our intuitive grasp of the principle of universal benevolence does not prove that it is a substantive normative truth, but we consider it makes that a reasonable hypothesis to hold, at least until a better explanation is offered.⁴⁵

There are thus three elements in the process of establishing that an intuition has the highest possible degree of reliability:

1. careful reflection leading to a conviction of self-evidence;
2. independent agreement of other careful thinkers; and
3. the absence of a plausible explanation of the intuition as the outcome of an evolutionary or other non-truth-tracking process.

If the third requirement were not met—if the intuition could be explained as the outcome of an evolutionary process—that would not show the intuition to be false, but it would cast some doubt on its reliability. The agreement of others would not put this doubt to rest, for this agreement

44. *The Analects of Confucius*, also known as *The Selected Sayings of Kongfuzi*, XV, 23; *Mahabharata*, Anusasana Parva 113.8; for Buddhism, see *Samyutta Nikaya* v. 353. Note that the Golden Rule is not merely advocating reciprocity. It tells us to do unto others as we would have them do unto us, whether or not they actually do treat us as well as we treat them. The words of the *Mahabharata* are especially clear on the distinction between self-interest and concern for others: “One should not behave towards others in a way which is disagreeable to oneself. This is the essence of morality. All other activities are due to selfish desire” (Anusasana Parva 113.8). For a comprehensive list of such ideas in many different texts and civilizations, see Howard Terry, *Golden and Silver Rules of Humanity*, 5th ed. (West Conshohocken, PA: Infinity, 2011).

45. In correspondence on an earlier version of this article, Tersman agreed that, as far as Sidgwick’s principle of rational benevolence is concerned, “we presently don’t have a fully satisfactory and well-established (evolutionary) debunking explanation of it.” He indicated that such a debunking explanation might, in time, emerge. We accept that this is possible. Tersman added that the line of argument we have developed in this article is “vulnerable to possible falsification by future empirical results and empirical theorizing.” We agree, but we also think Tersman is correct when he goes on to say that although this vulnerability to future possible falsification should motivate some humility, on the part of both skeptics and nonskeptics, about their conclusions, it is not a drawback or fault. On the contrary, as Tersman puts it, “more philosophers should try to articulate their positions in a way that makes them vulnerable in this way” (e-mail to the authors, July 25, 2011). Our judgment as to how well grounded a moral principle is should be sensitive to our best understanding of how we have come to accept that principle.

could be explained by the fact that the others share the same biological nature. This would raise the possibility that, in thinking that the intuition is self-evident, we are deceiving ourselves. Because the intuition plausibly could arise as part of our evolved nature, it would be, as Street argues, a coincidence if it happened to also be true. Coincidences do sometimes happen, but if an intuition that met the first two requirements but not the third were to clash with an intuition that met all three, we would have a ground for preferring the intuition for which there was no evolutionary explanation.

On the other hand, if an intuition does not meet the first two requirements, the fact that it meets the third would not help it. The ideal of celibacy serves as an example. Celibacy seems likely to diminish, rather than enhance, reproductive fitness, although if it brings sufficient power or prestige, the benefits that the celibate might confer on his or her kin conceivably could outweigh the loss of direct descendants. The widespread support for celibacy during many centuries of the Christian era suggests that our moral ideas are not always responses to evolutionary pressures, but in the absence of some specific religious beliefs, few regard celibacy as an ideal, and certainly not as a self-evident one.⁴⁶

We have argued that Sidgwick's axiom of universal benevolence passes this test, but we are not claiming that it is the only principle to do so. Other principles, including deontological principles, might be equally impartial—for instance, the principle that lying is wrong, whether one is lying to strangers or to members of one's own community. Ethical principles of respect for human rights might also be thought to be impartial in the same way, but to be fully impartial, they would need to be freed from any specific association with members of our species and instead to be reformulated as rights that are possessed by all beings with certain capacities or characteristics.

As we mentioned earlier, the principle that the good of one individual is of no more importance, from the point of view of the universe, than the good of any other tells us nothing about what this good may be. The principle of universal benevolence needs a theory of well-being, or else it is empty of content. Sidgwick is, of course, aware of this. That is why in *The Methods of Ethics* his chapter on philosophical intuitionism, in which he defends the axioms, is followed by a chapter in which he argues that the ultimate good is pleasure and the absence of pain. Whether Sidgwick succeeds in that endeavor is something we leave open, as we will also leave the question of what kind of method could be used to determine what

46. But why, one might ask, does celibacy not meet with the opprobrium of incest or homosexuality, given that it seems equally likely to diminish reproductive success? We do not know the answer, but could it be that for most humans it is less tempting, and therefore people are less in need of dissuasion?

is ultimately good. Kahane claims that most plausible theories of well-being, including hedonism, are obvious candidates for evolutionary debunking.⁴⁷ Our primary aim in this article is to show that partial reasons can be debunked and that, whatever the ultimate good may be, we have overriding reasons to aim at it impartially, so in response to Kahane's contention we will limit ourselves to pointing out that if no theory of well-being or intrinsic value were immune to a debunking explanation, this would show only that no theory could be preferred over others on the ground that it alone cannot be debunked. It could not show that no theory of well-being is true.

V. THE DUALISM OF PRACTICAL REASON

Now that we have prepared the ground, it is not difficult to see the implications of our argument for Sidgwick's "profoundest problem." It is, Sidgwick believes, "in accordance with common sense to recognize—as Butler does—that the calm desire for my 'good on the whole' is authoritative; and therefore carries with it implicitly a rational dictate to aim at this end."⁴⁸ This may indeed be in accordance with common sense, but here common sense seems likely to have been formed by the evolutionary influences we have been discussing. Since the claim that egoism is rational clashes with the principle of universal benevolence, we have precisely the situation described in the previous section, and we have grounds for supporting the intuition for which there is no evolutionary explanation rather than the one for which there is an evolutionary explanation. If the rationality of egoism can thus be put in doubt, we can tentatively conclude that all reasons for action are impartial, and the dualism that led Sidgwick to fear "an ultimate and fundamental contradiction in our apparent intuitions of what is Reasonable in conduct" can, at least on the level of rationality, be dissolved.⁴⁹

This may seem too paradoxical to take seriously. Utilitarians face a similar issue when defending the impartiality of utilitarianism. Sidgwick's response was to point out that although utilitarianism is impartial at the level of theory, in practice there are various factors that limit the extent to which we should try to act impartially, including our greater knowledge of how to bring about our own happiness—which is of course a part of the general happiness—as compared with the difficulty of knowing what will increase the happiness of strangers. Sidgwick also notes that we are

47. Kahane, "Evolutionary Debunking Arguments," 120; for a thoughtful response to debunking explanations of hedonism, see Knut Skarsaune, "Darwin and Moral Realism: Survival of the Iffiest," *Philosophical Studies* 152 (2011): 229–43. We owe this reference to Derek Parfit.

48. Sidgwick, *Methods of Ethics*, 112.

49. *Ibid.*, 508.

better able to increase the happiness of others when we are happy ourselves.⁵⁰ In a similar manner, the common view that it is rational to act self-interestedly may gain plausibility because acting in one's own interest, broadly conceived, is often in harmony with doing what is in the best interests of all. Nevertheless, this harmony is far from complete. In a world with a wide gulf between rich and poor, and many opportunities for the rich to help the poor, impartiality remains highly demanding for the rich.

Some of the remaining air of paradox around the idea that all reasons for action are impartial stems from the assumption that a reason for action must provide the person for whom it is a reason with a motivation for acting. Denying the rationality of egoism leaves reason detached from our strongest sources of motivation, namely, our desires to further our own interests and those of our family. If, however, we follow Thomas Nagel, Thomas Scanlon, Jonathan Dancy, and Derek Parfit in distinguishing normative reasons from motivating reasons, the paradoxical nature of our claim is reduced.⁵¹ On this view, normative reasons are independent of our present desires, wants, and beliefs. A normative reason can be a motivating reason when we act for this reason. But we may also have a motivating reason without having a normative reason. Parfit gives the example of someone who acted in order to get revenge. We may say, "His reason was to get revenge, but that was no reason to do what he did."⁵² A discussion of motivating reasons is, Parfit believes, relevant to why people act as they do but not to how they ought to act. The distinction is important because it allows for a conception of practical reason that is free of Hume's assumption that reasons for action must be based on desires. We can have normative reasons for action, irrespective of whether we like them, agree with them, or desire to act in accordance with them.

Given Parfit's insistence on the normative rather than the psychological nature of practical reason, our argument suggests that he could have gone further and rejected what he refers to as personal and partial reasons. Why then does Parfit accept the validity of personal and partial reasons, rather than say that they are very common *motivating* reasons but—as with the desire for revenge—not *normative* reasons? One possibility is that, like so many contemporary moral philosophers, he accepts the model of reflective equilibrium made popular by John Rawls, and this leads him to be reluctant to reject too many of our common moral judgments. But Parfit interprets reflective equilibrium widely, so that the process of reaching an equilibrium takes into account both scientific theo-

50. *Ibid.*, 431.

51. See Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986), chap. 8; T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998), chap. 1; Jonathan Dancy, *Practical Reality* (Oxford: Clarendon, 2000), chap. 1; Parfit, *On What Matters*, 1:37.

52. Parfit, *On What Matters*, 1:37.

ries and normative theories. As he puts it, "When we try to achieve what Rawls calls reflective equilibrium, we should appeal to all of our beliefs, including our intuitive beliefs about the wrongness of some kinds of act."⁵³ Among the scientific theories to be taken into account is evolutionary theory, along with the argument that it undermines the credibility of some of our most widely shared moral intuitions. Parfit, in particular, is well aware of this, for he stated it with his usual clarity in *Reasons and Persons*:

if some attitude has an evolutionary explanation, this fact is neutral. It neither supports nor undermines the claim that this attitude is justified. But there is one exception. It may be claimed that, since we all have this attitude, this is a ground for thinking it justified. *This* claim is undermined by the evolutionary explanation. Since there is this explanation, we would all have this attitude even if it was not justified; so the fact that we have this attitude cannot be a reason for thinking it justified. Whether it is justified is an open question, waiting to be answered.⁵⁴

Parfit, and other proponents of reflective equilibrium, widely interpreted, could therefore draw on evolutionary theory, as well as on Sidgwick's normative arguments, in order to reject many widely shared moral intuitions, while retaining the principle of universal benevolence. Although those who make use of reflective equilibrium in normative and applied ethics typically assume that they should try to achieve an equilibrium between a plausible normative theory and most, or at least many, of our commonly accepted moral judgments, there is no need for them to make this

53. *Ibid.*, 367. Street, too, refers to the "widespread consensus that the method of reflective equilibrium, broadly understood, is our sole means of proceeding in ethics" ("Darwinian Dilemma," 124). Street's acceptance of this model may have led her to neglect the possibility of defending moral realism that we have adopted, namely, that of accepting that many of our common moral intuitions are false, while defending at least one fundamental principle that we reach by the use of our reason. If we can grasp some moral truths by the use of our reason, while we hold others only because their acceptance enhanced our evolutionary fitness, it would be a mistake to assume that the best normative view is the one that holds these two kinds of beliefs in reflective equilibrium. Incidentally, despite the now-common view that reflective equilibrium, widely interpreted, is the inescapable method of justification in ethics, it is possible to interpret Sidgwick as a foundationalist who does not appeal to all of our beliefs but rather to a limited number of self-evident axioms. For discussion, see Peter Singer, "Sidgwick and Reflective Equilibrium," *Monist* 58 (1974): 490–517; Steven Sverdlik, "Sidgwick's Methodology," *Journal of the History of Philosophy* 23 (1985): 537–53; David Brink, "Objectivity and Dialectical Methods in Ethics," *Inquiry* 42 (1999): 200–210; Anthony Skelton, "Henry Sidgwick's Moral Epistemology," *Journal of the History of Philosophy* 48 (2010): 491–519; Roger Crisp, "Sidgwick and the Boundaries of Intuitionism," in *Ethical Intuitionism: Re-evaluations*, ed. P. Stratton-Lake (New York: Oxford University Press, 2002), 56–75; and Phillips, *Sidgwickian Ethics*, chap. 3.

54. Derek Parfit, *Reasons and Persons* (Oxford: Clarendon, 1986), 186; we owe this reference to Kahane, "Evolutionary Debunking Arguments," 110.

assumption. They could reject the commonly held view that it is rational to do what is in one's own interests (even though people may have strong motivating reasons to act in this way) and accept that when one of two possible acts would make things go impartially better, that is what we have decisive normative reason to do.