Robert Shaver is Professor of Philosophy at the University of
Manitoba.

Sidgwick on
Moral Motivation

Robert Shaver

In the first chapter I spoke of actions that we judge to be right and what ought to be done as being “reasonable,” or “rational” … and I contrasted the motive to action supplied by the recognition of such reasonableness with “non-rational” desires and inclinations. This manner of speaking is employed by writers of different schools, and seems in accordance with the common view and language …. For we commonly think that wrong conduct is essentially irrational, and can be shown to be so by argument; and though we do not conceive that it is by reason alone that men are influenced to act rightly, we still hold that appeals to the reason are an essential part of all moral persuasion …. On the other hand it is widely maintained that, as Hume says, “Reason, meaning the judgment of truth and falsehood, can never of itself be any motive to the Will”; and that the motive to action is in all cases some Non-rational Desire, including under this term the impulses to action given by present pleasure and pain. It seems desirable to examine with some care the grounds of this contention before we proceed any further (ME 23).¹

So opens Methods I.III. Sidgwick holds that moral judgments are claims about what it is reasonable to do. In Methods I.I he

writes of “rational procedure,” “rational precepts,” the “end of reasonable human action,” and “reasonable conduct.” When we fail to do what we see to be right, we are acting on “irrational springs of action” or “non-rational inclinations” (ME 1-6). He also holds that these judgments about what it is reasonable to do can motivate. He must, then, respond to Hume’s argument that reason cannot motivate. Most of Methods I.III is this response.

I shall clarify Sidgwick’s claim that moral judgments are judgments about what it is reasonable to do (I). I then give two interpretations of his view of moral motivation (II-III): insofar as one is rational, moral beliefs cause desires; or insofar as one is rational, one has a standing desire to act rightly, and moral beliefs combine with this desire to motivate. These views differ from the accounts often given by anti-Humeans that rely on entities such as “besires”—accounts which there is no need to adopt (IV). I then give Sidgwick’s argument against Hume, defending the view that judgments about what is reasonable can motivate (V), and I reply to the following objections (VI): the argument is question-begging; Hume can escape the argument by endorsing a standing-desire model of motivation; anti-Humean views cannot account for obvious slack between belief and motivation; and Sidgwick’s strategy wrongly puts the burden on Hume. I close with an objection I do not answer: Hume might explain moral motivation as the noncognitivist does.

Two preliminaries. First, Sidgwick can be seen as offering a version of Michael Smith’s “moral problem”:

(i) Moral judgments are beliefs about what it is reasonable to do.
(ii) Moral judgments sometimes motivate.
(iii) Beliefs alone do not motivate.²


Sidgwick holds (i) and (ii), rejecting (iii) [on one reading of (iii)]. Hume, on Sidgwick’s and the received view, holds (ii) and (iii), rejecting (i). I shall sometimes make reference to the debate over Smith as a way to illuminate Sidgwick.

Second, when I write of moral judgments’ motivating, I intend either of two possibilities: the agent takes the moral judgment to be backed by his all-things-considered judgment about what he has most reason to do, and so desires to perform, and performs, the action enjoined by the moral judgment; or the agent does not take the moral judgment to be backed by his all-things-considered judgment, and so has merely a desire (or “some motivation”) to perform the action, which is defeated by desires backed by other reasons. Sidgwick does not hold that moral judgments need always be the winning motivation, even in rational agents.

I

The topic is Hume’s objection that if moral judgments are claims about what is reasonable, they cannot motivate. But first the position Sidgwick is defending should be sketched.

Following Smith, one can distinguish rationalism as a conceptual position (that when I make a moral judgment, I am making a claim about what it is reasonable to do) from rationalism as a substantive position (that there are requirements of rationality that match up with moral requirements).³ By “rationalism” below, I mean rationalism as a conceptual position.

Sidgwick does not give an explicit argument for rationalism. He finds rationalism “in accordance with the common view and language .... For we commonly think that wrong conduct is essentially irrational, and can be shown to be so by argument” (ME 23). Rationalism is the position we need

³ Smith, Problem pp. 63-5.
This might seem far-fetched. “Irrational” has a range of meanings. On a narrow account, irrationality requires internal inconsistency: I believe what I think I have decisive reason to reject, or I do what I think I have decisive reason not to do. It is unlikely that wrong conduct always involves this sort of irrationality. On a wide account, irrationality requires only that one is subject to rational criticism. Sidgwick holds that I can be subject to rational criticism without internal inconsistency. For example, in rejecting the view that there are only hypothetical imperatives, he writes that “[w]e do not all look with simple indifference on a man who declines to take the right means to attain his happiness, on no other ground than that he does not care about happiness. Most men would regard such a refusal as irrational” (ME 7). After arguing that various axioms, such as the supervenience of the moral on the non-moral, pass his tests for highest certainty, Sidgwick comments that “I regard the apprehension … of these abstract truths, as the permanent basis of the common conviction that the fundamental precepts of morality are essentially reasonable” (ME 383). Reasonableness lies in passing epistemic tests for clarity, self-evidence, and consistency within and between people; internal consistency is not enough. Similarly, Sidgwick often describes customs, biases, prejudices, and privileges as being “irrational,” where this seems best glossed as “against the balance of reasons.” All this suggests that for Sidgwick, wrong conduct is irrational in the sense of being subject to rational criticism. Indeed, Sidgwick distinguishes between moral judgments and claims such as “the food is disagreeable” by noting that “the peculiar emotion of moral approbation is … inseparably bound up with the conviction … that the conduct approved is ‘really’ right—i.e., that it cannot, without error, be disapproved by any other mind” (ME 27). “Reason” rather than “moral sense” is appropriate because “the term Sense suggests a capacity for feelings which may vary from A to B without either being in error … and it appears to me fundamentally important to avoid this suggestion” (ME 34). Rationalism can be thought of as claiming that moral judgments, unlike simple reports or expressions of feeling, are subject to rational criticism, although the criticism need not consist in showing an internal inconsistency.

II

Sidgwick seems to claim that belief and rationality are sufficient for moral motivation. For example, “[w]hen I speak of the cognition or judgment that ‘X ought to be done’ … as a ‘dictate’ or ‘precept’ of reason to the persons to whom it relates, I imply that in rational beings as such this cognition gives an impulse or motive to action” (ME 34). These cogni-

---

4 For a useful discussion, see T. M. Scanlon, What We Owe to Each Other (Cambridge, MA: Belknap, 1998) pp. 25-33.
5 Sidgwick sometimes uses “irrational” in this way (e.g., ME 5, 23, 24, 37, 236, BB 160, PD 670, LE 23, 29).
6 For this account, see Derek Parfit, Reasons and Persons (Oxford: Oxford University Press, 1984) p. 119.
7 ME 269-70, PPE 390, OHE 34-5, E 297, CE 47, DEP 154. In the first edition of the Methods, Sidgwick writes that “one meaning of ‘irrational’ as applied to conduct is ‘inconsistent.’” This might seem to favour the narrow account of irrationality. But Sidgwick goes on to distinguish two “grades” of inconsistency: “Our impulses to action may be such as not to conflict, and yet not harmonized or systematized: or they may be actually opposed and conflicting.” As an example of the latter, Sidgwick suggests “when desiring an end we decline to take the necessary means … or if aiming generally at a kind of objects or results, we shew ourselves arbitrarily indifferent to a particular individual or instance.” As an example of lack of system, Sidgwick suggests maxims “that cannot be bound together and firmly concatenated by means of some one fundamental principle” (ME [1] 25-6). “Inconsistent,” then, turns out to mean something like “subject to rational criticism.”
Robert Shaver

Sidgwick on Moral Motivation

tions “in so far as they relate to conduct on which any one is deliberating ... are accompanied by a certain impulse to do the acts recognised as right” (ME 77). If Sidgwick did not hold that belief and rationality are sufficient for motivation, there would be no need for an argument against Hume.

Sidgwick does not, however, hold that beliefs motivate directly, without desires. Cognitions give or are accompanied by “motives” or “impulses.” When reviewing Methods I.III, Sidgwick writes of “the impulse that prompts us to obey the dictates of Reason” (ME 39). Later he claims that “we expect that the judgments of moral goodness or badness, passed either by the agent himself or by others, will—by the fresh motive which they supply on the side of virtue—have an immediate practical effect in causing actions to be at least externally virtuous” (ME 426-7). A desire is a “felt impulse ... to actions tending to the realisation of what is desired” (ME 43n2). “Motives” are “desires that we feel for some of the foreseen consequences of our acts” (ME 60; also 202-3, 362-3). Throughout, Sidgwick uses “motive,” “impulse,” and “desire” interchangeably (e.g., ME 43ff, 362ff).9

This picture is reinforced by passages elsewhere.

First, after giving Hume’s account of how reason can interact with desire, Sidgwick objects that “the account just given of the influence of the intellect on desire and volition is not exhaustive” (ME 25). This implies that for Sidgwick, as for Hume, reason has a role in motivation by influencing desire; the difference between Sidgwick and Hume is not over whether desires are present, but rather over how reason influences desire. Similarly, Sidgwick describes Hume as holding that “the motive to action is in all cases some Non-rational Desire” (ME 23). The “non-rational” qualification suggests that Sidgwick allows other, “rational” desires to play a role in motivation. (For Sidgwick, an “irrational desire” is “a desire directed to an action which we at the same time judge to be bad or incompatible with our highest good: e.g. desire of the pain of another human being, in the case of malevolent impulse—of escape from a post of duty and danger, in the case of a cowardly impulse” [LE 23; also LE 29, 31]. It is a desire “impelling us to volitions opposed to our deliberative judgments” [ME 24]. Non-rational desires are desires that produce actions “without our having judged such actions to be either right or wrong” [ME 24]. Rational desires are presumably desires directed to actions which we judge to be good. One can choose “without rational choice—i.e. not as good” [LE 31].)

Second, Sidgwick accepts Green’s description of him as holding that “the opposition between Desire, Reason, and Will” lies “between different desires, of which reason however ... supplies the object to the one, while some irrational appetite is the source of the other; the will being the arbiter which determines the action according to the rational or irrational desire” (LE 22). Rational action is action on the basis of (rational) desire.

Third, against Leslie Stephen, Sidgwick claims that “the feelings that normally cause action are not pleasures and pains as such, but desires and aversions.”10 It is, for exam-

9 Usually for Sidgwick a desire is something “felt.” It is a state “in which we are conscious of an impulse to get out of the present state into a future one” (FT 98). But Sidgwick does not always require an occurrent feeling: “not unfrequently during long processes of work for remote ends, the desire of the end, while remaining sufficiently strong to supply the requisite impulse to action, ceases to have a perceptible character as feeling; we only infer its presence from the actions that it stimulates, and from the satisfaction that follows on the attainment of some intermediate end which has no significance for us except as a step towards the ultimate end” (FT 99). The feeling of desire “may temporarily cease to be present in ... consciousness,” though it “will normally recur at brief intervals during the process” (ME 55).

10 Sidgwick does not explain the “normally.” Perhaps he is thinking of reflex actions, which may be caused by a feeling that is not a desire or aversion. A similarly passing comment, restricted to voluntary actions, occurs in the Methods: “I shall use the term Motive to denote the desires of particular results, believed to be attainable as consequences of our voluntary acts, by which desires we are stimulated to will those acts” (ME 363).
ple, aversion that “normally accompany[ies] the moral judgment that an act is wrong.” Stephen holds that if this aversion is not an aversion to the pain of remorse or pain caused to others, the moral judgment “cannot, any more than any other operation of the intellect, move the will, apart from emotion.” Sidgwick replies that “[t]his seems to me a rather subtle psychological question; but at any rate I should say that in the case of moral judgments I am frequently not conscious of any emotion other than what is implied in the mere consciousness that the judgment carries with it a motive” (RS 573). Sidgwick holds that desires are (at least normally) needed for action and that moral judgments carry desires with them.

Fourth, those who hold that beliefs motivate directly usually give motivation by duty as their example. But Sidgwick seems uninterested in taking motivation by duty as motivation without desire. In his argument against the view that I desire only my own pleasure, he notes as a counter-example “the love of virtue for its own sake, or desire to do what is right as such” (ME 52). Elsewhere he sometimes characterises this motivation as “the desire ... of doing what is right as such, realising duty or virtue for duty’s or virtue’s sake,” “a pure desire to act rightly” (ME 204, 206). He even characterises Kantian moral motivation as “pure Regard for Moral Law, Desire to do Right as Right” (ME [1] 346).11

III

It is important to see precisely what Sidgwick rejects in Hume, given that he agrees that desires are needed for motivation.

11 The corresponding passage in the seventh edition reads “pure regard to duty or choice of Right as Right” (ME 366; also GE 185). There Sidgwick also describes Kantian motivation as “disinterested obedience to Law as such, without the intervention of sensible impulses” (ME 513).

Smith’s characterisation of “the Humean theory” claims that motivating reasons are constituted by belief-desire pairs and that “for any belief and desire pair that we imagine, we can always imagine someone having the desire but lacking the belief, and vice versa.”12 This does not rule out holding that beliefs (sometimes) cause desires with the pair then motivating. Indeed, this is Smith’s own account of moral motivation.13

Smith’s characterisation does not fit Hume. For Hume would lose his main anti-rationalist argument were he to agree that moral beliefs sometimes cause desires (or, more precisely, “motivating passions”), without the contribution of pre-existing desires. He could not argue that rationalists fail to capture the connection between moral judgment and motivation.14 Nor could he argue that, since reason is “inactive,” it cannot conflict with the passions.15

Sidgwick must refute Hume’s anti-rationalist argument. To do so, it is sufficient to argue that moral beliefs some-
times cause desires, without the contribution of pre-existing desires. (Call this the “causal” model.) Since Sidgwick takes moral beliefs to be beliefs about justifying reasons, he would describe one whose moral beliefs fail to cause the relevant desire as imperfectly rational; rationality lies (in part) in the functioning of this causal mechanism. His claim, then, is that moral beliefs and rationality are sufficient for motivation. Hume must deny the possibility of the causal mechanism. I take this, rather than the conceptual separability of particular beliefs and desires, to be the issue.

Sidgwick may be tempted by a different position. He might claim that the needed desire is produced by the belief because, insofar as one is rational, one has a standing desire to do what one takes to be right. (Call this the “standing-desire” model.) Some passages suggest the standing-desire model: we have “a determination to pursue whatever conduct may be shown by argument to be reasonable,” a “genuine impulse to conform to the right rules of conduct,” and “a disinterested aversion to unreason” (ME 5, 101, 447). The passages concerning a “desire to do what is right as such” cited earlier also suggest this model (although they are typically offered as descriptions of Stoic or Kantian views, rather than as Sidgwick’s own account of ordinary moral motivation). But Sidgwick’s wording usually suggests the causal model: that the cognition “gives an impulse or motive” or is “accompanied by” an impulse or “carries with it a motive” or “supplies” a “fresh motive” suggests that the impulse is not already present. More importantly, if Sidgwick held the standing-desire model, there would be little issue with Hume over whether reason can motivate. Hume and Sidgwick would agree that motivation requires a pre-existing desire. The issue between them would instead concern whether reason just supplies “ideas of actual or possible facts,” since Sidgwick thinks reason supplies moral judgments and that moral judgments do not convey ideas of actual or possible facts (ME 25). This is a possible reading of Methods I.III, but it makes Sidgwick’s concern with motivation at the start of the chapter misleading.

18 For the causal model, see Parfit, “Reasons and Motivation” 105; Radcliffe 101-3, 117; Rachel Cohon, “Hume and Humeanism in Ethics,” Pacific Philosophical Quarterly 69, 1988, 103, 105-6, 108; Jonathan Dancy, Moral Reasons (Oxford: Blackwell, 1993) pp. 8-9, 12. Stroud may suggest it but is probably better read as identifying desires with a disposition that need not be “before the mind.” See Stroud, pp. 168, 179.

19 I put the point in terms of perfect rationality (or “one feature of rationality”) because one might argue that when my belief about what I have reason to do is false, it is sometimes more rational for me to have a desire opposed to this belief than for me to have a desire consistent with this belief. (For this point directed against Smith, see Nomy Arpaly, “On Acting Rationally Against One’s Best Judgment,” Ethics 110, 2000, 493-501.) One can agree but still claim (as Arpaly admits) that having a desire opposed to my belief about what I have justifying reason to do is a falling off from perfect rationality.

16 For the controversy over Hume’s actual position, see Radcliffe, and Cohon, Virtue chs. 2-3. In the context of the dispute with Sidgwick, Hume might admit the causal mechanism but deny that this mechanism is (even partially) constitutive of rationality, perhaps by invoking an account of rationality that is inapplicable to the fit between beliefs and desires. On such an account, weakness of will would not count as irrationality. This, however, makes his disagreement with the rationalist semantic—what the rationalist, and indeed most of us, consider rationality is sufficient, when combined with beliefs about reasonable action, to motivate.

17 For a standing-desire model as a rationalist reply to Hume, see Jonathan Harrison, Hume’s Moral Epistemology (Oxford: Clarendon, 1976) pp. 13-15. In his initial exposition, C. D. Broad seems to attribute the model to Sidgwick, although his later gloss is more ambiguous: compare “[h]uman beings have an impulse or desire to do what they judge to be right …. It is … one motive among others” with “in most human beings the belief that a certain course of action is right … is pro tanto a motive for doing it …. We are perfectly familiar with this motive” (C. D. Broad, Five Types of Ethical Theory [London: Routledge and Kegan Paul, 1930] pp. 146, 179).

20 This is hardly decisive, however. As one referee notes, a particular moral belief such as “one ought to be brave” could be said to give or supply a fresh motive in the sense that it gives direction to the standing desire to do what one takes to be right.

21 Hastings Rashdall, after quoting much of Methods I.III, also suggests both the causal and standing-desire models, but like Sidgwick seems to favour the former. He writes, for example, that the “recognition of the thing as right is capable of producing an impulse to the doing of it”; the desire to do what we see to be our duty “may be created by the Reason which recognizes the rightness”; “an act may be done … from a desire which is created solely by our conviction that
The causal model may represent a change in position. In the first edition of the *Methods*, in reply to the contention that "Reason cannot be a spring of action, as it must always be Feeling that stimulates the Will," Sidgwick claims that no one is competent or really concerned to maintain that the apprehension of duty is a state of consciousness which occurs without any emotional element. Hence we need not ask whether a mere cognition can act upon the Will and prompt to action. It is enough if it be granted that there exists in all moral agents as such a permanent desire ... to do what is right or reasonable because it is such: so that when our practical reason recognises any course of conduct as right, this desire immediately impels us ... towards such conduct ... We may assume then as generally admitted that the recognition of any action as reasonable is attended with a certain desire or impulse to do it: and that in this sense the Reason may be affirmed to be a spring of action (ME [1] 22, 27, 28).22

This explains moral motivation by finding it unnecessary to hold that moral judgments sometimes motivate by themselves. Moral agents have a standing desire to do what they take to be right. If this is satisfactory, the argument of the later editions, given in V below, is unnecessary. Not surprisingly, it is absent from the first edition; a rough version of it appears, and the quoted passage from the first edition disappears, in the second edition (ME [2] 25-31).

Sidgwick had good reason to abandon the standing-desire model as it is stated in the first edition. For if I lack a certain course is intrinsically right or reasonable.” Against Kant, he requires desires for motivation. See Rashdall, *The Theory of Good and Evil*, second ed. (London: Oxford University Press, 1924) v. i, pp. 104, 106, 121, 106n2 (also 140-1). 22 J. B. Schneewind also notes that Sidgwick seems to have changed his position (though Schneewind says very little about what the post-first-edition position is). See Schneewind, *Sidgwick’s Ethics and Victorian Moral Philosophy* (Oxford: Clarendon, 1977) p. 235; for Schneewind’s terse comment on the later view, see p. 207.

The standing desire, then on this model I might, with perfect rationality by my own lights, make a moral judgment without the desire to act on it, just as I might believe there are anchovies on a pizza without being rationally compelled to desire the pizza. This could not appeal, since for Sidgwick (in all editions), if I think x ought to be done, then it follows that I think I have a justifying reason for doing x. I could not then, with perfect rationality by my own lights, make the judgment without the desire. That would be to hold that I have a justifying reason to do x without desiring to do x—a failure of perfect rationality.23 (Perhaps this explains why Sidgwick cut the first-edition passage.)

Sidgwick could, however, adopt an augmented version of the standing-desire model. He could add that it is part of the concept of a rational agent that she have a desire to do what is right.24 For a rationalist, this is very plausible: since thoughts about what is right are thoughts about what is rational, the claim becomes that it is part of the concept of a rational agent that she have a desire to do what is rational.

One worry about combining rationalism and a standing-desire model is raised by J. L. Mackie. Against Jonathan Harrison’s suggestion that rationalists can answer Hume by adopting a standing-desire model, Mackie objects that

[j]t is not merely that it is linguistically odd to use words like “right” and “wrong” with no prescriptive force—to say, for example, “X is right and Y is wrong, but of

23 For a similar argument, see Smith, *Problem* pp. 62, 150-2.

24 One might read the first edition as saying just this by supposing that Sidgwick makes it part of the concept of a moral agent that she desires to do what is reasonable. Sidgwick, however, goes on to appeal to “experience” to show the presence of the desire to do what is reasonable, suggesting that he takes the desire to be conceptually independent of the concept of a moral agent (ME [1] 27)). Alternatively, one might note that Sidgwick sometimes runs together the cognition that x ought to be done with the impulse or desire to do x (e.g., ME 39), perhaps because he thinks the cognition causes a desire, as in the causal model. On this interpretation, the desire is not, as the long quotation seems to suggest, independent of the cognition.
Robert Shaver

Since the standing-desire model makes claims only about motivation, Mackie must, in arguing that Harrison loses prescriptivity, be identifying prescriptive force with motivating force. But rationalists should deny this identification. Consider an analogy: say that given a conditional and its consequent, I infer the antecedent. It is up to me whether I do so in the sense that I may or may not be motivated to do so. But it is not up to me whether I do so in the sense that doing so would be a mistake. Rationalists, then, can answer Mackie by accepting that motivation is contingent on a standing desire without accepting that justification is contingent on a standing desire.

I conclude that Sidgwick opposes Hume’s view of moral motivation by adopting the causal model, but that he could also do so by adopting a standing-desire model, augmented by the claim that it is part of the concept of a rational agent that she have a desire to do what is right. If Sidgwick accepted this standing-desire model, his argument against Hume’s view of moral motivation in Methods I.III would be unnecessary.

IV

Today, unitary states that are both belief- and desire-like are the usual opponents of the Humean theory. (Mackie, quoted above, also sees them as the historical opponent.) But Sidgwick need not attack Hume by subscribing to such unHumean states.

One reason for favouring unHumean states is phenomenological: one might take oneself to be motivated to act on moral judgments without noticing any desire. The Humean can avoid this by rejecting a phenomenological conception of desire. More importantly, this does not seem to be Sidgwick’s thought: when he discusses the phenomenology of acting on moral judgments, he never stresses the absence of desires, and usually notes their presence.

Sidgwick on Moral Motivation


26 Stephen Darwall gives a different argument against Sidgwick’s holding a standing-desire model. He thinks Sidgwick’s reason for rejecting non-normative analyses of moral judgments is that moral judgments, unlike non-normative judgments, are “intrinsically action guiding.” If so, Sidgwick cannot hold that the action-guidingness comes from a standing desire rather than from the moral judgment itself. However, as noted, Sidgwick’s rejection of non-normative analyses relies on pointing out specific problems encountered by specific analyses—for example, apparently opposed judgments turn out to be consistent—rather than on any appeal to intrinsic action-guidingness. See Darwall, “Intuitionism and the Motivation Problem,” in Philip Stratton-Lake ed., Ethical Intuitionism (Oxford: Clarendon, 2002) pp. 269-70.
Since the phenomenological argument for unHumean states is the most prominent, more should be said. Consider David McNaughton’s statement of the argument:

Reflection on what it is like to realize that one is morally required to act in a certain way will ... support [the] contention that a moral belief can be sufficient to motivate an agent .... Once we are aware of such a requirement our choice of action seems to be constrained by that recognition .... In particular, the claims that morality makes on us appear to be quite independent of our desires—they may even conflict with what we want. On the belief-desire theory no sense can be made of this conception of a moral demand. Moral ... considerations can only provide an agent with reason to act if he has appropriate desires .... [Thus Humeans] must claim that the apparent authority of moral demands is an illusion.30

This is unconvincing for two reasons.

First, unHumean states are not needed to avoid dependence on antecedent desires; the causal model allows for motivation without antecedent desires. Kant himself often seems to hold that respect is a feeling caused by a moral belief; the feeling causes a desire (or “interest”) which then motivates.31

Second, Sidgwick can, without embracing unHumean states, say that moral demands are independent of my desires in that moral demands concern what it is reasonable to do, and what it is reasonable to do need not depend on what I desire. (My desires are exclusively relevant only if the best normative theory is the instrumental theory.) Moral consid-
erations can provide desire-independent justifying reasons—desire-independent constraints. These can be and feel authoritative in that they limit the options I can choose while remaining rational.32

Jonathan Dancy gives a different argument. We should favour unHumean states over the causal model since on the causal model,

sometimes beliefs need the independent contribution of desire, and sometimes they don’t. This seems to mean that there are two sorts of beliefs, Humean and non-Humean. But that cannot be quite right, because the very same beliefs which on one occasion need an independent non-cognitive supplement if there is to be an action, can elsewhere figure in a state which needs no such supplement at all. Beliefs are not carved up into two sorts, those that can motivate alone and those that need some help. It is the same sort of belief all the time. But in that case how is it that the extra provided by a desire can on other occasions be provided by a belief?

Dancy concludes that the anti-Humean should not concede that desires are ever needed: the causal model “will be unable to retain its insistence that purely cognitive motivation is possible if it once yields to the pressure to admit Humean motivation as a possibility.”33

The causal model can be defended. On it, one need not agree that the same beliefs sometimes need an independent desire to motivate and sometimes do not. Nor need one agree that the extra is provided by more beliefs. The beliefs that do not need independent desires, but instead cause desires, are beliefs about what it is reasonable to do. They cause desires not everywhere, but only insofar as one exhibit-

30 McNaughton pp. 48-9.
32 For this second strategy against McNaughton, see W. K. Frankena’s (much earlier) “Obligation and Motivation in Recent Moral Philosophy,” in Frankena, Perspectives on Morality, ed. K. E. Goodpaster (Notre Dame: University of Notre Dame, 1976) pp. 51-60.
33 Dancy p. 21.
its that feature of rationality that requires consistency between one’s beliefs about what it is reasonable to do and one’s desires; rationality is the extra. The beliefs that need independent desires to motivate—such as “there are anchovies on the pizza”—do not cause desires because the rationality of the believer does not combine with their content to cause a desire. More generally, although it might be mysterious to carve beliefs into two sorts according to their effects *simpliciter*, there is nothing mysterious about carving beliefs into two sorts according to how their different contents react in the presence of another property (rationality).34

I conclude that there is no case for adopting unHumean states, and so no reason to reject Sidgwick’s anti-Humean position on the ground that it does not involve such states.

34 One referee rightly notes that in places, Sidgwick does not mention rationality; beliefs seem to do all the work. (i) In “The Morality of Strife,” Sidgwick writes “[s]uppose, for instance, that every one who is liable to drink too much had clearly present to his mind, in the moment of temptation, the full amount of harm that his insobriety was doing to his bodily health, his reputation, his means of providing for those dependent on him; some, no doubt, would drink all the same, but the great majority of those not yet in bondage to the unnatural craving would draw back. Suppose, again, that any one who is wronging a neighbour saw, as clearly as any impartial judge or friend would see, the violation of right that he is committing; surely only a thoroughly bad man would persist in his wrong-doing” (MS 49). Sidgwick does not mention rationality here, but one can read the reference to “bondage” as suggesting that the belief motivates only given the absence of the practical irrationality that addiction brings. This does not cover the case of the man wronging his neighbour. But Sidgwick adds that, thoroughly bad men aside, one wrongs “from a mixture of intellectual muddle with passionate impulse or selfish negligence” (MS 49). This is compatible with thinking that wrong-doing is the result of either mistaken moral beliefs or the sort of practical irrationality suggested by “passionate impulse” and “negligence.” (ii) In “Unreasonable Action” (discussed briefly in VI), Sidgwick again emphasises that practical irrationality stems from one’s beliefs’ not being clearly present to the mind. But again, this is consistent with holding that rationality plays a role; irrationality lies in not keeping one’s belief clearly present to the mind. Moreover, Sidgwick does claim that irrational action is possible even when the belief is clearly present to the mind; here something beyond the belief is needed as an explanation (UA 139-42).

Sidgwick’s reply to Hume begins by noting that “[e]very one, I suppose, has had experience of what is meant by the conflict of non-rational or irrational desires with reason.” For example, we feel “bodily appetite prompting us to indulgences which we judge to be imprudent, and anger prompting us to acts which we disapprove as unjust or unkind” (ME 23-4). Sometimes, our judgment that we ought not to act imprudently or unjustly stops us from acting imprudently or unjustly. On the face of it, this is an example of reason motivating.

Sidgwick notes that there is a Humean tactic for handling this potential counter-example. The Humean understands the moral judgment as describing facts about either consequences of action or means to ends that combine with pre-existing desires. Thus

(1) I desire to hit Jones.
(2) I hold that I ought not to hit Jones.
Therefore I do not hit Jones.

becomes (say)

(1) I desire to hit Jones.
(2)* I believe that not hitting Jones is a necessary means to satisfying my desire to use Jones’s swimming pool this summer.
(3) I desire to use Jones’s swimming pool more than I desire to hit Jones.
Therefore I do not hit Jones.

On the Humean view, “the experience which is commonly described as a ‘conflict of desire with reason’ is ... more properly conceived as merely a conflict among desires and
aversions; the sole function of reason being to bring before the mind ideas of actual or possible facts, which modify in the manner above described the resultant force of our various impulses” (ME 25).

Sidgwick rejects this account of the conflict. Moral or prudential judgments “have some ... influence on volition,” but “cannot legitimately be interpreted as judgments respecting the present or future existence of human feelings or facts of the sensible world.” Sidgwick’s argument for this negative claim consists in “showing the inadequacy of all attempts to explain” moral or prudential judgments that do not take them to be different from “all notions representing facts of physical or psychical experience” (ME 25). Hence Sidgwick argues that, say, “one ought to tell the truth” cannot be analysed by any such representative claim. He proceeds by examining candidate analyses such as “truth-telling is an efficient means to some end” or “truth-telling excites a feeling of approbation in me” or “not truth-telling will be penalised by God.” He rejects each. The instrumentalist analysis leaves out evaluations of ends; the subjectivist analysis leaves out disagreement; the voluntarist analysis leaves out questions of the rightness of the penaliser (ME 26-31).

The argument thus runs as follows: Humean analyses of (2) in the manner of (2)' are unsatisfactory; hence the Humean cannot account for some cases in which reason seems to motivate; hence we should reject the claim that reason does not motivate; hence we can keep the claim that moral judgments are judgments about what is reasonable that sometimes motivate.\(^3^7\)

VI

I shall consider four objections.

(A) There is an obvious objection. The Humean could refuse to admit Sidgwick’s example of a motivating moral judgment as a case of reason motivating. For he would reject Sidgwick’s identification of moral judgments with judgments of reasonableness. If this identification were rejected, the Humean could agree that moral judgments can motivate without accepting that reason can motivate.

The objection can be put differently. Sidgwick’s motive for rejecting the view that reason cannot motivate is his wish to maintain that moral judgments are judgments of reasonableness (given that he also believes that moral judgments motivate). It begs the question, in defence of the belief that moral judgments are judgments of reasonableness, to assume that moral judgments are judgments of reasonableness.

Sidgwick has a reply. For he can run his argument without supposing that moral judgments are judgments of reasonableness. Rather than offer moral judgments as motivating, he can offer judgments of reasonableness as motivating. ([2] could be replaced by “I hold that hitting Jones is unreasonable.”) He can then argue against the Humean strategy of handling potential counter-examples to the claim that reason cannot motivate by analysing judgments of reasonableness as specifying means to desired ends or facts that cause further desires. The same arguments that block this analysis for moral judgments are applicable to judgments of reasonableness.

Robert Shaver

Moreover, one can say this without adding that rational agents must have a desire to do what is right. One can allow, as the rationalist does not, that I can fail to be motivated by my moral beliefs without any irrationality by my own lights.

However, if a Humean moves to this model to avoid analysis of moral judgments, the objection that rationalists cannot capture moral motivation is lost. For on this model, thoughts such as “I ought not to hit Jones” can motivate, whatever gloss one gives to them, by combining with a desire to do what one thinks is morally right. The analysis of the moral judgment is irrelevant to this explanation of motivation (as it is not, for example, on an account according to which moral judgments describe or express one’s desires.) The rationalist merely gives a particular gloss—thinking “I ought not to hit Jones” is thinking “hitting Jones is unreasonable.” And if there is no objection to the rationalist gloss, there seems no reason to deny that failure to be motivated by moral beliefs manifests irrationality.

(C) Anti-Humeans face a standard objection: two people can have the same beliefs yet different motivation, and so their beliefs alone cannot be what motivate them. For Sidgwick can allow, however, that the same beliefs can


Sidgwick on Moral Motivation

ableness. If “x is reasonable” is analysed as “x is an efficient means to some end,” we cannot claim that ends are reasonable; if “x is reasonable” is analysed as “x excites approbation in me,” parties disagreeing over reasonableness do not really disagree; if “x is reasonable” is analysed as “acts other than x will be penalised by God,” we cannot say that some divine penalties are reasonable. Sidgwick can conclude that since the Humean strategy of handling potential counter-examples fails, there is no objection to thinking that reason can motivate, and so no objection to thinking that if moral judgments are claims about what is reasonable, moral judgments can motivate.

(B) A Humean might object that no analysis of moral judgments is needed to explain Sidgwick’s cases. Hume argues that we confuse “calm desires” such as “the general appetite to good, and aversion to evil, consider’d merely as such” with the calm sensation of reasoning. For Hume, the “general appetite to good” is a desire for pleasure, and so moral judgments would need to be connected to pleasure, perhaps by the sort of analysis Sidgwick rejects. But the strategy of combining a standing desire with a moral judgment shows how one could avoid any analysis: moral motivation can be explained by citing a desire to do what one takes to be right. For example,

(1) I desire to hit Jones.
(2) I hold that I ought not to hit Jones.
(3)’ I desire to do what I believe to be morally right more than I desire to hit Jones.

Therefore I do not hit Jones.

For the problem, see Dancy pp. 12, 22; Cohon, “Hume and Humeanism” 111-12. Dancy’s reply is that a belief can motivate in one case but not in another, without forcing one to say that a desire must explain this difference and so must form part of the explanation in both cases. For (i) the absence of factors that would have blocked a causal chain are not typically seen as parts of the chain, and so need not enter into an explanation; and (ii) the explanation need not cite a desire at all—Dancy (like Cohon) gives, as examples of blocking factors, carelessness, inattention, despair, drunkenness, neurophysiological disorders, and depression (Dancy pp. 22-6). (i) is controversial: one might see this as a merely pragmatic point concerning what we wish to focus on, and so as leaving open the possibility that desires still play a key role (even if in the background). (ii) can be read as making the point Smith and Sidgwick make: moral beliefs motivate practically rational agents. (ii) does not, however, justify thinking that beliefs can suffice for motivation without desire, since it is plausible to think that depression and the like prevent motivation by preventing the desires usually caused by moral beliefs.
cause different motivations. One agent might be irrational, and so the belief either fails to produce a desire or produces a desire that is defeated by other desires. Since Sidgwick’s claim is that what suffices for desire are moral beliefs and that feature of rationality that requires consistency between my beliefs about what it is reasonable to do and my desires, the explanation of the failure of motivation is not restricted to the belief. Sidgwick thinks that in most cases of irrationality, “seductive feeling” makes me temporarily forget my practical judgment, or makes me think it inapplicable to the case at hand, or makes me pay attention to only some of the relevant reasons; in the latter two cases, I can be “obscurely conscious” that I am making a mistake. These are all effects on my beliefs. But Sidgwick does allow that in rarer cases, my practical judgment is “clearly present in [my] consciousness” yet I am not motivated (UA 139-42).

(D) Sidgwick starts by assuming that there are apparent cases of reason motivating. My judgment that I ought not to act out of anger sometimes prevents me from so acting. Sidgwick proceeds by rejecting Humean redescriptions of these cases and concludes that reason does indeed motivate. But later generations, raised on a diet of Hume, will be chary about accepting any apparent cases of reason motivating. What Sidgwick gives as a Humean redescription will strike many now as the natural description: there is “merely a conflict among desires and aversions.”

One reply is that what present-day Humeans insist on is that desires are needed for motivation. Unlike Hume, they do not insist that beliefs cannot cause desires. (It is hard

41 Smith objects to anti-Humeans that “[i]t is a commonplace, a fact of ordinary moral experience, that when agents suffer from weakness of the will they may stare the facts that used to move them square in the face, appreciate them in all their glory, and yet still not be moved by them” (Problem p. 123). Sidgwick agrees that this is possible—as it seems McNaughton does not—but he doubts that, in most cases of weakness of will, the facts are appreciated in all their glory. For McNaughton, see pp. 128-31.

even to find much discussion of whether beliefs can cause desires.) If so, present-day Humeans need not resist Sidgwick’s description of cases of reason motivating as cases in which beliefs cause desires.

A different reply is that, if necessary, Sidgwick can run his argument without the starting assumption that there are apparent cases of reason motivating. Sidgwick can reject, as he does, particular redescriptions of what the moral judgment conveys; he can add that he and Hume agree that moral judgments motivate; he can conclude that since the motivating power of moral judgments does not come from (say) their telling us means to satisfy our desires, their motivating power seems to come simply from their being the sort of judgment they are: they are judgments that cause desires insofar as one exhibits that feature of rationality that requires consistency between one’s beliefs about what it is reasonable to do and one’s desires. Moral judgments also seem to be judgments about what is reasonable, given the way in which we argue over them. If so, moral judgments are examples of reason motivating.

This version of Sidgwick’s argument makes clear, however, a final objection. One might explain the motivating power of moral judgments as the noncognitivist does. Sidgwick does not consider a noncognitivist analysis, although his rejection of a subjectivist analysis points the way: “truth-telling excites a feeling of approbation” can become an expression rather than (or in addition to) a description of attitude, so that parties disagreeing over truth-telling are indeed disagreeing, by expressing opposed attitudes. There is no need for Hume to concede that reason can motivate, since the apparent cases of reason defeating desire can be understood, as the noncognitivist understands them, as a clash of desires.

One reply would be to construct objections to noncogni-
tivism from Sidgwick. But whether Sidgwick would object to noncognitivism is not so clear. This may be surprising, given Sidgwick’s status in histories of ethics as a “non-naturalist” or “intuitionist” against whom the noncognitivists developed their alternative. But Sidgwick’s commitment is to rationalism rather than to cognitivism or “descriptivism.” Much depends on the success of noncognitivism at capturing the features of moral argument Sidgwick wields against other analyses. Ambitious noncognitivists, such as Hare and Gibbard, might give Sidgwick all he wants.42

Introducing noncognitivism suggests a more comprehensive gloss on Sidgwick’s overall argument. Sidgwick and Hume agree that moral judgments motivate. They might motivate (a) by describing facts that combine with pre-existing desires; (b) by describing desires; (c) by being desires or the like; (d) by combining with a pre-existing desire to do what one takes to be right; (e) by causing desires, without the need for pre-existing desires; (f) by expressing desires. Sidgwick’s arguments against analyses of moral judgments, if successful, rule out (a) and (b). (c) is unnecessary given the possibility of (e). (d) does not raise a problem for rationalism. (e) is Sidgwick’s account. Whether Sidgwick could instead adopt (f) depends on the possibility of rationalist noncognitivism.43

42 For the same thought, see Scanlon pp. 59, 64.
43 Thanks to Shane Huson, Tim Schroeder, Sergio Tenenbaum, two anonymous referees, and, especially, Joyce Jenkins for comments on earlier drafts.