Hume's Reasons
Aaron Zimmerman
Hume Studies Volume 33, Number 2, (2007) pp. 211-256.

Hume Studies
Volume 33, Number 2, November 2007, pp. 211–256

Hume's Reasons

AARON ZIMMERMAN

Abstract: Hume's claim that reason is a slave to the passions involves both a causal thesis: reason cannot cause action without the aid of the passions, and an evaluative thesis: it is improper to evaluate our actions in terms of their reasonableness. On my reading, Hume motivates his causal thesis by arguing that accurate representation is the function of reason, where a faculty of this kind cannot produce action on its own. (The interpretation helps vindicate Hume of the common charge that he "begs the question" against his opponents.) But Hume's causal thesis does not entail his evaluative thesis, and his commitment to the latter is incredibly thin. According to Hume's positive theory, our evaluative judgments originate in reason integrated with sympathy or humanity. And, I argue, the resulting view depicts us as having substantive, non-instrumental reasons to fulfill our obligations to both prudence and morality.

A great deal of Hume's *Treatise* argues for limitations on the scope and power of reason. Our belief in persistent physical objects that are external to and distinct from our minds, our belief that causes necessitate their effects, and our belief that our selves are something other than a collection of perceptions are all argued to be the products of imagination instead. But what is perhaps Hume's most famous attempt to cut reason down to size emerges from his discussion of human action and its psychological origins. "Reason," Hume says, "is and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them" (T 2.3.3.4; SBN 415).[1]

Aaron Zimmerman is Associate Professor in the Department of Philosophy, 5707 South Hall, University of California, Santa Barbara, California 93106-3090. E-mail: azimmerman@philosophy.ucsb.edu.

Hume's slogan is suggestive, but it is couched in highly metaphorical language. Even if we think of the mind as divided into modules and allow reason and the various passions among their number, we must still make sense of one module's being enslaved to some others. What, then, is the literal view of practical reason to which Hume gives voice when he speaks of the subjugation of reason to passion? What led him to describe the relation in such derisive terms?

In one sense, the answer to this question is fairly uncontroversial. Reason is a slave to the passions because reason cannot *cause* actions on its own. No matter how well (or poorly) the faculty of reason functions, its operations must be augmented with the outputs of a distinct faculty if action is to result. ("Reason alone can never produce any action, or give rise to volition" (T 2.3.3.3; SBN 414).) And because reason is in this sense "motivationally impotent," reason alone cannot combat the outputs of our passions—faculties that perhaps can yield volitions and actions without the assistance of reason. ("The same faculty [reason] is incapable of preventing volition, or of disputing the preference with any passion or emotion.") Thus, Hume concludes, it cannot be "the understanding" alone that enables us to resist our self-destructive and licentious appetites; calm passions play an essential role. Prudence and morality are not the products of the unaided exercise of reason, and "[t]he greatest part of moral philosophy, antient and modern," is mistaken in asserting otherwise (T 2.3.3.1; SBN 413).[2]

Of course, a man will not be able to figure out what is good for him on the whole without the help of reason, as he must use his understanding to determine the causal impact his present actions will have on his future wellbeing.[3] But something more than this is necessary if he is to actually act prudently. He must have some concern for his future good—a concern that will owe its existence, at least partially, to fear of future suffering, a calm desire for his good "as such," or the output of some distinct "passionate" faculty. Similarly, although Hume is quite clear that a man will be unable to fully distinguish right from wrong if he cannot infer a person's motives from her actions and gauge the impact these motives will have on the pleasure, happiness or well-being of those who display them and those they affect, something more than knowledge of motives, actions, and their utilities is necessary if a man is to be motivated to act morally. He must be endowed with "sympathy" or "humanity" if he is to feel disapprobation toward vice and approbation toward virtue. In the absence of sentiments or feelings of these kinds (however faint) his cold knowledge of the destructive effects of his prospective actions will not dissuade him from their performance. Insofar as we think of our efforts to act prudently and morally as a battle between our passions on the one hand and unaided reason on the other, we are mistaken.[4]

Still, while Hume's account of the role that reason plays in the production of action is relatively clear, obscurity enters when we try to: (i) determine his reasons for endorsing the view, and (ii) explain why he seems to draw such shockingly

skeptical conclusions from it. First, exactly why does Hume think "that reason alone can never be a motive to any action of the will"? At least two arguments for this claim are given at T 2.3.3, and significant disagreement exists as to both their content and quality. Second, even supposing that Hume does have a reasonably good argument for the motivational impotence of reason, how is the purported fact that reason cannot *cause* action without the help of some passion supposed to support Hume's subsequent claim that self-defeating, imprudent, and immoral actions cannot be properly *evaluated* as unreasonable (T 2.3.3.6; SBN 415–6)? Consider a prudent man's preference for his long-term good that is generated by reason acting in concert with the normal operations of sympathy or humanity. We might consider, for instance, as Hume would describe the case, a man who understands that smoking will increase the likelihood of future illness, understands that his future self would hate being ill, and sufficiently identifies—that is, empathizes and sympathizes—with that future self to prefer the lesser immediate pain of abstinence to the greater future pain he thinks he would otherwise experience. If the man subsequently kicks the habit, why can't his preference be properly awarded the honorific "reasonable"? And why can't we appropriately call "unreasonable" someone's indifference to his own future wellbeing, especially when this lack of concern accounts for the continued indulgence of appetites he knows are self-destructive? Does the purported fact that unaided reason could not move a man to abandon smoking imply, all on its own, that it is not unreasonable for him to smoke? What more must be assumed?

My aim in what follows is to answer these questions, and in so doing formulate a charitable, well-grounded reading of Hume's view of practical reason. First, I will try to show that a degree of interpretative unkindness accounts for the common charge—leveled by Christine Korsgaard, Elijah Millgram, and others—that Hume's arguments for the motivational impotence of reason beg the question against his opponents.[5] On Korsgaard's reading, it is only because Hume assumes that we must use something above and beyond reason to figure out how we ought to act that he concludes that reason alone cannot produce action or do battle with the passions. As Korsgaard sometimes puts it, Hume's "motivational skepticism" is grounded in a question-begging "content skepticism."[6] Millgram joins Korsgaard in claiming that Hume's first argument for the motivational impotence of reason assumes content skepticism of this kind. But according to Millgram, Hume's second (more radical) argument is premised on Hume's now discredited view that passions, emotions, and desires are wholly "non-representational" entities.

Admittedly, Hume thinks that a person must sympathize (or have sympathized) with others if she is to acquire genuinely moral beliefs, and he probably thinks a person must have (or must have had) some aversion to her future suffering if she is to make genuinely prudential judgments. And it is also true that Hume rejects the (neo-Aristotelian) suggestion that concern for one's future and sympathetic

response to the joys and pains of others are themselves aspects of reason. Nevertheless—contra Korsgaard—these are not premises in the better of the two arguments Hume offers at T 2.3.3.[7] Moreover, while Hume does indeed think that our passions are not themselves representations, but are only "annexed" to ideas that represent their objects,[8] this controversial claim is not, as Millgram suggests, an essential premise in the second of Hume's two arguments for the enslavement of reason to passion. Instead, a more charitable construal would have Hume leaning on the far more plausible proposition that truth or accurate representation is not the *function* of anger, joy, fear, sympathy, or any of the other so-called "passionate" faculties he concludes must always be implicated in action.

On the reading I will defend, Hume grounds his conclusions at T 2.3.3 in an important asymmetry between the faculty that generates demonstrative knowledge and causal belief on the one hand, and those faculties that generate volition and action on the other. For Hume reasons, on both empirical and methodological grounds, that the faculty responsible for demonstrative and causal reasoning is distinguished from our other mental faculties by its properly functioning so as to represent what is real or true, whereas the passions are individuated via their quite different ends. Thus, I will argue, if we are to dispute Hume's demotion of the faculty of reason to the position of servant, we must reject truth as the aim of theoretical reason, demonstrate how truth really is the aim of action and volition, or adopt a non-functionalist taxonomy of the mind's faculties. None of Hume's premises beg the question against his opponents in any obvious way, and none can be jettisoned without argument.

This brings us to the essay's second task: to explain why Hume seemingly reasons from the motivational impotence of reason to the conclusion that no action or passion can be properly criticized as "unreasonable." Surprisingly, Korsgaard and Millgram don't criticize Hume for drawing this inference. Instead, they take the implication for granted, and then assume—understandably enough—that if Hume claims that no action is properly said to be any more "reasonable" than any other, he must think that we either lack all reason to act or that we lack all practical reasons save those that are instrumental in nature. For instance, Korsgaard describes Hume as thinking, "If the general desire for the good does not remain predominant, not only the motive, but the reason, for doing what will conduce to one's greater good, disappears" ("Skepticism," 15). And Millgram writes, "Hume is a skeptic, not an instrumentalist, if nothing can count as a reason for action, then the considerations adduced as instrumental reasons cannot count as reasons for action either" (80). In contrast, I think a more balanced reading of the text reveals that Hume's claim that no choice, passion or action can be properly called "unreasonable" results from a fairly superficial error regarding the commitments the common speaker incurs in using the expression. Admittedly, there may be someone out there who in calling a man's action or project "unreasonable" means to imply that it must

have been generated by a malfunction in that man's faculty of reason, and Hume would take issue with those who use "unreasonable" intending this implication; for as we will see, on the view Hume defends in the *Treatise* and *Enquiries*, some of the actions we call "unreasonable" do not arise from breakdowns in the faculty best identified with reason, but from impairments in dissociable human capacities for empathy, sympathy, and resolution. But, I will argue, most of us do not use "unreasonable" intending any such implication. Of course, interpretive charity has its limits, and it is right to point out that Hume draws an inference to which he is not entitled. But there are good reasons not to emphasize Hume's skeptical pronouncements—or to join Korsgaard and Millgram in reading Hume as inferring from them that we have no good reason to do anything—as his substantive (descriptive) account of how we actually come to criticize actions as weak-willed, imprudent and immoral neither imperils the truth or propriety of these forms of criticism, nor grants allegations of irresolution or instrumental irrationality an epistemic status denied judgments of imprudence and immorality. To be clear: I will not be arguing that Korsgaard and Millgram are ignorant or grossly mistaken about Hume's view of evaluative and normative judgment. Instead, I will suggest that their allegations of wholesale skepticism result from a failure to adequately take Hume's views on these matters into account when interpreting his view of our reasons to act in a prudent and moral fashion.

Let us say that *full skepticism* regarding practical reason is the view that we have no more reason to pursue one course of action than any another. In partial contrast, *instrumentalism* about practical reason is the view that we only have reason to pursue a given course of action if it is something we want or intend, or the means to the achievement of something we want or intend.[9] Thus, if Hume thought that we often have good, on-balance reasons to act, he was not a full skeptic about practical reason; and if he thought someone's having good reasons to pursue a given course of action does not always depend on the desires, ends or preferences that agent actually has, then he should not be accused of endorsing a wholly instrumentalist conception of practical reason. So, if we are going to evaluate Korsgaard's and Millgram's skeptical interpretations in a substantive manner we must ask: Did Hume really think that the ordinary man has no more reason to be prudent than imprudent? Are men who prefer their "lesser good" to their greater really supposed to lack such reasons?

The matter is complicated by flexibility in our ordinary use of "reasons" and by Hume's never giving us an explicit account of reasons for action. He relegates his claims to reason: the faculty, reasoning: its operations, and our classification and criticism of preferences and actions as "reasonable" and "unreasonable." So the texts are not as clear on the issue as one would like them to be. But what the texts do make clear, I will argue, is that Hume believed that we ought to be prudent and kind even when there is nothing we want that we can get by doing so. And

it is also relatively clear that Hume thought we are obliged to act in these ways because it would be vicious not to: *full stop*, where imprudence and cruelty owe their status as vices to the integrated operations of the understanding, empathy, and sympathy. And it is because I think a non-instrumentalist, non-skeptical view of reasons for action best fits Hume's view of the nature of virtue, vice, and their relation to obligation, that I think a charitable interpretation of the text as a whole would clear him of Korsgaard's and Millgram's allegations. On Hume's considered view, I will suggest, reason in concert with sympathy or humanity gives us non-instrumental reasons to act prudently and morally—we have these reasons by way of what Hume follows Shaftesbury and Hutcheson in calling a "moral sense."[10]

My plan is as follows. In §1–§2 I defend the teleological reading of Hume's claim that reason is enslaved to the passions. Then, in §3, I outline the basic account of prudence Hume provides at *Treatise* 2.3.3–4 and go on to argue that it does not provide Hume's complete account of practical reason. In §4 I begin to fill out Hume's view of reason's practical or action-guiding role, by describing his theory of how we come to criticize and condemn imprudence. Finally, in §5–§6 I argue that Hume is not best interpreted as providing an instrumentalist or skeptical account of reasons for action, and explain how he thinks an imprudent man's grasp of non-instrumental reasons might play a role in moving him to act prudently.

1. Hume's Reasoning

The outline of Hume's initial argument for reason's impotence is fairly transparent. Reason's possible functions are enumerated, and the generation of action is not included among them. The subsequent paragraph then argues that since reason alone cannot cause action on its own, it cannot inhibit the passions from doing so (T 2.3.3.2–4; SBN 413–5).

Argument #1

(1) The understanding can do no more than: (a) demonstrate relations between ideas so as to generate conceptual knowledge, and (b) employ causal or probabilistic reasoning so as to generate beliefs concerning matters of empirical fact.

(2) Demonstration of the relations that hold between our ideas cannot generate action on its own.

(3) Reasoning that establishes matters of empirical fact (even when combined with wholly demonstrative reasoning) cannot generate action on its own.

Therefore,

(4) The understanding cannot generate action on its own.

Therefore,

(5) The unaided understanding cannot subvert, control, or overcome the passions.

As I've said, Korsgaard and Millgram are right to charge this first argument with "begging the question" against Hume's opponents, for of the argument's three premises, only the second would be regarded as uncontroversial by his intended audience. (Hume is therefore right to assume that the denial of this second premise "scarce will be asserted" (T 2.3.3.2; SBN 413).) For instance, as Hume well knew, the rationalist Samuel Clarke had argued that a priori understanding is all that is needed to establish that gratitude is a more fitting or appropriate response to kindness than is indifference—and that one therefore ought to respond to kindness with gratitude rather than indifference. But even Clarke would not have suggested that this realization could generate an act of gratitude all on its own. Rather, empirical reasoning must be used to establish that one has received an act of kindness in some particular case, and that returning the favor (or writing a thank-you note) here and now would constitute an act of gratitude.[11] Even Kant, who would go on to argue that pure a priori reason can give us knowledge of the general maxims under which we are obligated to live, admitted that the operations of this faculty "require a power of judgment sharpened through experience, partly to distinguish in which cases they have their application, and partly to obtain access for them to the will of the human being . . . who is not so easily capable of making it effective *in concreto* in his course of life." (Ak 4:389)[12]

But what of the argument's two other premises? At this point in the *Treatise*, Hume has in hand a number of conclusions regarding the role the understanding can and cannot play in abstract reasoning and the drawing of causal inferences. But he has not yet shown that reason does not have some distinct, decidedly practical function. Thus, premise (1) seems to beg the question against those who would include as part of reason Aristotelian *phronesis* or practical wisdom: a faculty responsible for generating prudential and moral knowledge along with the (perhaps constitutively connected) emotional and appetitive responses characteristic of a good or virtuous person. Moreover, by assuming premise (3), Hume fails to engage (what contemporary philosophers would call) "cognitivist internalist" opponents who think: (a) that some combination of demonstrative and causal reasoning can lead a man to judge that, say, refusing to eat any fruits and vegetables would be imprudent, or that refusing to help a child in need would be immoral, and (b) that beliefs of this evaluative or normative kind need no assistance from the passions to motivate men to eat properly and help those in dire straits. For at this

point in the *Treatise*, Hume has done little to argue: (i) that judgments of prudence and morality cannot be generated by the understanding alone, or (ii) that those prudential and moral judgments that are generated by the unaided operation of the understanding cannot motivate action on their own. It isn't until later in the *Treatise* and *Enquiries*, that Hume considers these issues with the care and attention they deserve.[13]

Nevertheless, Hume's case for the subjugation of reason to passion does not end at T 2.3.3.3–4. For after acknowledging that the conclusion he draws from his premises is apt to shock his audience, Hume sets out to "confirm" his initial claims with some further "considerations" (T 2.3.3.4; SBN 415). Reason, Hume argues, must be understood in terms of representation. But the appetites and emotions driving us to imprudence and immorality—the passions that reason is supposed to overcome in the service of right action—are not representational entities of the relevant kind.

> A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five foot high. 'Tis impossible, therefore, that this passion can be oppos'd by, or be contradictory to truth and reason; since this contradiction consists in the disagreement of ideas, consider'd as copies, with those objects, which they represent. (T 2.3.3.5; SBN 415)[14]

It is because the passions have no "representative qualities" that "render them copies" capable of truth or falsity, accuracy or inaccuracy, that they cannot be quashed or overcome by reason.

The passage is admittedly obscure, but there are a number of fairly uncontroversial desiderata that any adequate interpretation of it must satisfy. First, the conclusion that Hume claims to be supporting at T 2.3.3.5 has to do with the understanding and the passions: faculties of mind that generate outputs from certain inputs. But Hume's argument for that claim is entirely focused on the supposed outputs of those faculties. (Ideas, Hume here claims, are copies, and so are true if they accurately represent that of which they are copies, whereas passions, emotions and appetites are not.) An adequate reconstruction of the argument must include some hypothesis as to how Hume reasoned from the supposedly truth-evaluable nature of the understanding's issue to the inefficacy of the faculty itself. Second, reason's motivational impotence must have something to do with the purported fact that our passions cannot be evaluated as true or false. Even if some other asymmetry between the processes that give rise to causal belief and

those that give rise to action could be found in Hume's work, the argument given here at T 2.3.3.5—and then elaborated at T 3.1.1.9—insures this much. Third, the conclusion of the argument must be that reason alone cannot cause action, and so cannot cause prudent or moral action in the face of imprudent or immoral inclination. No doubt, in the very next paragraph Hume goes on to draw the decidedly normative conclusion that passions, actions, and preferences cannot be properly called "unreasonable." But if we are to take Hume at his word, T 2.3.3.5 is supposed to "confirm" the conclusions of his first argument, which is clearly an argument for reason's motivational impotence; and without supplementation, the claim that motives and actions cannot be properly criticized as unreasonable neither entails nor is entailed by the proposition that unaided reason cannot produce actions and volitions. Fourth, and last, at T 3.1.1.8 Hume tells us that of all the arguments with which he has "proved" that reason is "perfectly inert" it is the argument from the non-representational nature of the passions that bears restating in a more "conclusive" form. He continues,

> Reason is the discovery of truth or falsehood. Truth or falsehood consists in an agreement or disagreement either to the real relations, or to real existence and matter of fact. Whatever, therefore, is not susceptible of this agreement or disagreement, is incapable of being true or false, and can never be an object of our reason. Now 'tis evident our passions, volitions, and actions, are not susceptible of any such agreement or disagreement; being original facts and realities, compleat in themselves, and implying no reference to other passions, volitions, and actions. 'Tis impossible, therefore they can be pronounc'd either true or false, and be either contrary or conformable to reason. (T 3.1.1.9; SBN 458)

If it is to adequately capture Hume's mindset, an adequate rendering of T 2.3.3.5 must find within it the seeds of the more developed T 3.1.1.9.

We can therefore begin with a reconstruction of 2.3.3.5 that draws heavily on 3.1.1.9. (Asterisks are used to mark the argument's premises.)

Argument #2

(1*) The function of the understanding is the generation of true or accurate beliefs. ("Reason is the discovery of truth or falsehood" [T 3.1.1.9; SBN 458].)

Therefore,

(2) The understanding can infix beliefs by inferring their truth or probable truth, and it can eliminate beliefs by inferring their falsity or probable falsity. For instance, reason can eliminate belief in the existence of objects

that do not exist, and undermine the false belief that a proposed action really is sufficient to achieve the agent's designed end. (T 2.3.3.6; SBN 415–6)

(3*) The powers of reason are delimited by its function—reason's operations end when it has achieved its function. ("Whatever, therefore, is not susceptible of this agreement or disagreement, is incapable of being true or false, and can never be an object of our reason" (T 3.1.1.9; SBN 458).)

(4*) Passions, volitions, and actions can be neither true nor false. ("A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification" (T 2.3.3.5; SBN 415); "Our passions, volitions, and actions, are not susceptible of any such agreement or disagreement [i.e., truth or falsity]" (T 3.1.1.8; SBN 458).)

Therefore,

(5) Reason cannot directly produce passions, volitions, or actions, as this would take it beyond its function. ("Reason alone can never produce any action or give rise to volition" (T 2.3.3.4; SBN 414).)

Therefore,

(6) Unaided reason cannot subvert, control, or overcome our passions except by undermining beliefs that are holding our motives in place or guiding them in the direction of the will. ("'Tis impossible, therefore they [i.e., passions, volitions, and actions] can be pronounc'd either true or false, and be either contrary or conformable to reason" (T 3.1.1.9; SBN 458); "Passions can be contrary to reason only so far as they are *accompany'd* with some judgment or opinion" [T 2.3.3.6; SBN 416].)

Therefore,

(7) "Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them" (T 2.3.3.4; SBN 415).

The argument is fairly cogent, and it rests on just three premises: (1*) true representation is the function of reason, (3*) a faculty cannot perform operations that lie outside its function, and (4*) passions, actions and volitions can be neither true nor false. Moreover, though premises (3*) and (4*) are not beyond legitimate dispute, it is really only the first of these three premises that might be thought to expose Hume to the charge of "begging the question."

First, with regard to premise (4*), even if we agree with Korsgaard and Millgram (and the vast majority of the contemporary philosophical community) that

desires and intentions are themselves representational in nature (and are not, as Hume claims, just "annexed to" metaphysically distinct ideas that represent that which is desired or intended), we might still insist that desires and intentions represent in a way that prevents them from being properly evaluated as true or false. My belief is only true if things are as I believe them to be, but my desire or intention is not properly said to be false merely because things are not as I want or intend them to be.[15]

Admittedly, things are less clear with regard to certain emotions. Perhaps, for instance, my fear of the spider crawling up my arm is inaccurate so long as the creature is not truly dangerous or fearsome. And it may even be, as Thomas Nagel suggests, that, "[s]ympathy is not, in general, just a feeling of discomfort produced by the recognition of distress in others, which in turn motivates one to relieve their distress. Rather, it is the pained awareness of their distress as something to be relieved."[16] But it would be odd to say that accurate representation of danger is the *function* of fear or that the *function* of sympathy is an accurate representation of exactly which instances of distress ought to be relieved. Instead, "mere" (unemotional) beliefs about what things really are dangerous and which instances of distress are truly undeserved are more reliable and so better suited to this task. Thus, Hume might argue, the faculty of fear aims at the avoidance of harm and danger, and only generates representations of what is dangerous or fearsome to accomplish this end; and sympathy aims at the kind of conditional benevolent action necessary for familial and social harmony, where judgments of desert play a similarly instrumental role in its realizing this function.

Moreover, it seems that Hume's argument would still go through if he allowed that the operations of reason can directly undermine certain passions or emotions, just so long as such undermining is restricted to a demonstration of the inaccuracy of distinctively *non-evaluative* aspects of an emotion's content. (We would here equate undermining a passion with what Hume considered rejection of the belief or judgment supporting that passion.) To keep with Nagel's example, on Hume's considered view, reason can supplant or control someone's sympathy by showing him that its object is not truly in distress or that the sufferer's distress is deserved because the outcome of a magistrate's justice rather than a private citizen's malevolence. What is in fact ruled out by any weakening of premise (4*) strong enough to do the job is an imprudent passion's having a distinctively evaluative content, where that passion might then be subverted by reason's showing its content false just as a belief might be undermined through reason's demonstration of its content's untruth. So, for example, it cannot be that the drunkard's thirst for excessive booze represents the late night rounds as good for him, and that reason can remove or control that appetite merely by demonstrating the greater benefits of stepping away from the bar. To think of our appetites, emotions, and passions in this way would be to see them as "aiming at" or functioning so as to represent the truth—

albeit the truth about what is good for us. And though this is something that Hume flat-footedly denies, his doing so is not tantamount to begging the question against his intended opponents. It is certainly not as bad as leaning on the assumption that our passions have no representational qualities whatsoever.[17]

Second, premise (3*) has the air of a truism. Of course, there is a sense in which a functionally characterized entity can exceed its proper function. (Though designed to measure, a ruler can be used to punish; though naturally suited to pumping blood, the heart can serve as an inexact metronome; and though a carburetor is supposed to deliver to its engine an explosive mixture of vaporized fuel and air, it can function as a rather bulky doorstop.) But Hume is engaged in the scientific study of the mind, and so is interested in the more restricted sense of "possibility" supplied by premise (3*). That is, all Hume need be claiming here is that even when a man's reason is performing as it should, and so fulfilling its proper function, its operations needn't result in the passions, actions, and volitions we deem reasonable, prudent, or moral. As it turns out, Hume concludes, the construction of the human mind is such that the operations of a distinct faculty must be conjoined with a properly functioning reason if right action is to result.

But what should we make of the argument's first premise? If Hume's first argument begs the question against the neo-Aristotelian in assuming that reason can do no more than establish relations between ideas and determine matters of fact, does not the second fail to engage the same opponent when it assumes that true belief—rather than, say, right action—is reason's function? Though this is a fair allegation, I think it can be answered by looking more carefully at T 2.3.3 and the material that precedes it, for we can there find the basis for a strong (if non-demonstrative) argument for premise (1*).

We can begin by noting that Hume's avowed aim in his philosophical work is a systematic psychology that will match, in explanatory power and elegance, the natural philosophies of Boyle and Newton. He therefore sees the need to develop a non-arbitrary taxonomy of the mind's contents and their principles of operation. Though he begins this task with thought rather than volition, he does not assume at the outset that the application of "reason" is limited to the former. Admittedly, he initially picks out the referent of "reason" as that faculty (whatever it is) that generates demonstrative knowledge and core empirical belief; but he then asks himself, with a decided ecumenicism, whether in addition to these ideas the mechanism so identified might produce actions and volitions as well. He treats this question as neither wholly a posteriori nor wholly a priori or stipulative in nature.[18] Instead, the extension of "reason" is to be expanded or restricted in light of a methodological principle—that of "explaining all effects from the simplest and fewest causes" (T Intro.8; SBN xvii)—alongside substantive investigation into what unites demonstration with probabilistic inference. And to this end, Hume considers several alternatives to the truth-directed conception of reason he winds

up endorsing, where he only ends up characterizing reason as the faculty aimed at truth because he thinks this classification is part of the best—most predictive, most explanatory—account of the mind's operations.

Which alternatives does Hume consider? First, in the *Treatise*'s first Book, Hume examines and rejects the claim that reason can be simply characterized as the faculty of representation. Of course, the mathematical beliefs that result from demonstrative reasoning and the probabilistic judgments that result from causal inference are both representations, so reason must be capable of representation. But when I daydream and intentionally think up stories, the fantasies and idle musings I entertain are also representations, and they cannot be properly classified as the products of the understanding. Instead, Hume claims, though both sorts of representation arise from the imagination, "reason" or "the understanding" picks out that part of the faculty responsible for demonstrative knowledge and causal judgment.

> [T]he word, *imagination*, is commonly us'd in two different senses; and tho' nothing be more contrary to true philosophy, than this inaccuracy, yet in the following reasonings I have often been oblig'd to fall into it. When I oppose the imagination to the memory, I mean the faculty, by which we form our fainter ideas. When I oppose it to reason, I mean the same faculty, excluding only our demonstrative and probable reasonings. (T 1.3.9n22; SBN 117–8)

"The understanding," Hume goes on to say, in contrast with "the fancy" refers to "the general and more establish'd properties of the imagination" (T 1.4.7.7; SBN 267; cf. T 1.4.4.1–2; SBN 225–6).

Second, Hume considers and rejects a claim that Kant would go on to endorse: that reason is best thought of as *the faculty of rules*. On a Kantian view of the mind, reason is that within us that represents (and/or legislates) laws dictating how we ought to think and act. Indeed, it is this conception of the understanding that Kant leaned on to motivate a scheme of classification on which true belief and right action are products of a single "unified" faculty. As Kant says,

> We have defined the understanding in various different ways: as spontaneity of knowledge (in distinction from the receptivity of sensibility), as a power of thought, as a faculty of concepts, or again of judgments. All these definitions, when they are adequately understood are identical. We may now characterize it as the *faculty of rules*. This distinguishing mark is more fruitful, and approximates more closely to its essential nature.[19]

But against this view, Hume has already argued that our core empirical beliefs are generated by causal inferences that are not in fact rule-guided. Though a full

reconstruction of Hume's arguments on this score would be impractical here, we can at least consider Hume's oft-repeated claim that causal reasoning must always contain at least one transition between an impression and an idea.[20] Does Hume allow that these operations might be performed by a faculty of rules? Well, to take just one case (albeit a central one), when Hume tries to explain how the imagination generates our belief in unperceived objects and phenomena, he doesn't appeal to any representational properties at all. Our belief in objects and events that are external to and distinct from our minds is instead explained by certain intrinsic properties: "strength" and "vivacity" (T 1.4.2.40–1; SBN 208), and certain non-representational relations that our impressions enter into: their substantial "constancy" and "coherence" with one another, where impressions are constant when they are qualitatively identical but numerically distinct, and coherent so long as they are regular in the order in which they appear before the mind but neither qualitatively nor numerically identical (T 1.4.2.19–20; SBN 195–7). My belief that the fire continues to burn as I repeatedly vacate and reoccupy a view of the hearth surely owes its existence to the understanding. But in core cases like this one, the operations of the understanding do not involve the rule-guided processing of representations, but are instead generated by certain non-representational properties of (and relations between) our sensations.

Of course, Hume does not think that all causal reasoning consists in direct transitions from impressions to ideas. The explicit application of his "rules by which to judge of causes and effects" constitutes reasoning, and it is sure to involve rule-guided transitions between our ideas themselves.[21] But Hume is clear that central cases of causal reasoning—where the "inferred" cause or effect is more obvious—do not involve the application of rules. Indeed, he ends the relevant section by expressing his doubt that the rules that he has posited are ever really needed.

> Here is all the LOGIC I think proper to employ in my reasoning; and perhaps even this was not very necessary, but might have been supply'd by the natural principles of our understanding. Our scholastic head-pieces and logicians shew no such superiority above the mere vulgar in their reason and ability, as to give us any inclination to imitate them in delivering a long system of rules and precepts to direct our judgment, in philosophy. (T 1.3.15.11; SBN 175)

The core causal beliefs of ordinary people are the result of the understanding, and the understanding's products will either be commensurate or contrary to reason, but the formation of these beliefs is not at all rule-guided.

> For we here find, that the understanding or imagination can draw inferences from past experience, without reflecting on it; much more without

> forming any principle concerning it, or reasoning upon that principle. (T 1.3.8.13; SBN 104)[22]

Of course, this isn't to deny that causal belief formation exhibits regularities and therein conforms to the laws a psychologist might use to describe and predict our beliefs about the unobserved. But being accurately describable by rules does not distinguish the operations of reason from the behavior of non-human animals and plants, or even clocks, solar systems, and other inorganic unities. Thus, on Hume's reckoning, if we limit the scope of reason to operations that are truly rule-*guided*, we must exclude core cases of causal inference; but if we allow as operations of reason everything that merely *conforms* to rules, we will find ourselves driven to an absurdly rationalistic view of nature.[23]

Finally, at T 2.3.3 Hume explicitly considers a phenomenological scheme for individuating the mind's faculties. And he admits that a psychology founded on immediate introspection would grant reason its own motivational force by attributing prudential concern and benevolence to the same "faculty . . . which judges of truth and falsehood . . . because their sensations are not evidently different" (T 2.3.3.8; SBN 417). His offhanded rejection of this view as incompatible with what can be discerned with a "strict philosophical eye" reflects his growing dissatisfaction with the explanatory tools of the phenomenologist—a dissatisfaction he has earned by trying and failing to use immediately "felt" qualities to distinguish several other psychological entities from one another.[24] Thus, Hume concludes, on grounds both methodological and substantive, that the crucial characteristic of reason, the mark that allows for the most explanatorily fruitful theory, is not a feeling, but a natural effect or purpose—accurate representation—that does not constitute the end of motivation.[25] Since it would violate simplicity to use a conjunctive function to individuate a single faculty—as we would were we to regard true belief and right action as the goal of reason—Hume announces that "reason" is the faculty responsible for discovering truth and falsehood.[26]

It seems, then, that Hume has a response to Korsgaard's request that we "attempt to arrive at a general notion of reason by discovering features or characteristics that theoretical and practical reason share [by appealing to] such characteristic features as universality, sufficiency, timelessness, impersonality, or authority" ("Skepticism," 17). Hume hasn't begged the question against this project, but instead argued on largely methodological grounds that mental faculties are best individuated by their ends, while arguing on largely empirical grounds that the features Korsgaard cites are not the ends of theoretical reason. Timelessness can be ruled out in the first way: though it is, perhaps, a feature of theoretical reason, it isn't its function. Authority is ruled out in the second: since our basic causal beliefs do not result from a tacit grasp of rules expressing our epistemic obligations, theoretical skepticism would result from identifying

reason as whatever it is within us that responds to (what we believe or know to be) our duties.

This reconstruction of Hume's argument for the subjugation of reason relies on a straightforwardly functionalist reading of, "Reason is the discovery of truth or falsehood" (T 3.1.1.9; SBN 458). And I can foresee resistance to it arising from the natural tendency to interpret this sentence in a different way, by reading Hume as equating reason with truth, or as embracing the obvious falsehood that beliefs must be true or accurate to be rational. But Hume's considered view is that empirical beliefs are rational when they "fit" one's experience to date, even if regularities in that experience turn out not to be universal. As Hume would go on to say, "The Indian prince, who refused to believe the first relations concerning the effects of frost, reasoned justly" (EHU 10.1.10; SBN 113). It is, Hume says, a mental entity's "susceptibility" to truth or falsehood that is necessary if it is to qualify as an object or product of reason. The relation or (in Hume's words) "reference" to truth that something must have if it is to issue from reason concerns the point of its production. As John Bricke points out, the distinction between reason and passion motivating Hume at T 2.3.3 prefigures Elizabeth Anscombe's metaphor of the differing "directions of fit" of belief and desire, though Hume is able to literalize the metaphor by making a claim about faculties rather than states of mind.[27]

I conclude my case for this reconstruction of Hume's reasoning at T 2.3.3 with evidence that Hume deployed the teleological conception of reason in at least two other contexts in the *Treatise*. First, at T 1.4.1 Hume claims that the rules we employ in demonstrative reasoning are "certain and infallible" and that all errors in demonstration arise from their misapplication, as when we miscalculate a sum and fail to correct the error. When we think of reason as the calculating (or demonstrating) faculty we will see that truth is its end or proper function—an end that it often fails to accomplish.

> Our reason must be considered as a kind of cause of which truth is the natural effect; but such-a-one as by the irruption of other causes, and by the inconstancy of our mental powers, may frequently be prevented. (T 1.4.1.1; SBN 180)[28]

Second, when describing our attribution of identity over time to plants, animals and particular artifacts, Hume claims that causal interactions between temporary entities that "conspire" toward a common end or purpose help explain our thinking of them as parts of a single enduring object (T 1.4.6.11–2; SBN 257). He then extends this account to the soul, arguing that relations of government and subordination between our differing impressions, ideas, characters, and dispositions encourages belief in the substantial unity of these disparate items (T 1.4.6.19; SBN 261). It cannot be denied that our concepts of enslavement and subordination are

intellectual cousins. Thus, it seems that when Hume came to examine the relation of reason to the passions he applied his teleological account of our belief in the soul to the parts or faculties of which that soul is composed. And the conception of reason to result drove Hume to declare reason passion's slave.

2. Psychology and Normativity

Naturalistic reconstructions of Hume's argument for reason's enslavement inevitably give rise to a difficult objection. Hume's language, it is protested, is not purely psychological, but decidedly normative in nature. Perhaps we can hear Hume's claim that reason and the passions can never "oppose" each other or "dispute for the government" of our actions as purely causal in tone, but Hume then goes on to say that the understanding can neither "justify nor condemn" a passion, and that, "properly speaking," even a predilection for the world's destruction cannot be called "unreasonable" (T 2.3.3.6–7; SBN 415–7). If, as I will argue, Hume thinks that we should not prefer evil things, and he thinks that (irrespective of our desires) we have substantial reasons not to prefer them, indeed, if those endowed with sympathy and humanity can be led to appreciate these reasons by simply taking up the requisite points of view, why does Hume refuse to vilify destructive preferences with—to use Kieran Setiya's helpful phrase—the language of reason? As Thomas Reid complained,

> To act reasonably, is a phrase no less common in all languages, than to judge reasonably. We immediately approve of a man's conduct, when it appears that he had a good reason for what he did. And every action we disapprove, we think unreasonable, or contrary to reason. A way of speaking so universal among men, common to the learned and unlearned in all nations, and in all languages must have a meaning. To suppose it to be words without meaning, is to treat, with undue contempt, the common sense of mankind.[29]

The question, then, is whether Hume *really* thought—as a nihilist might—that every action and volition is just as reasonable or unreasonable as every other. And the short answer is that though Hume did assert as much, his commitment to the claim was incredibly shallow. For it was only because Hume neither conducted nor even pretended to conduct a careful examination of ordinary language that he mistakenly assumed that something can only be properly labeled "unreasonable" if it consists in or arises from a malfunction in the faculty of reason. We are just as happy to call a man "irrational" when he fails to do what he knows he must to be happy as when he remains unmoved by what we take to be a compelling argument, and we rarely stop to think about whether these diverse phenomena

have a common cause.[30] And were we to reflect on the matter and accept Hume's argument that irrational beliefs and weak-willed, imprudent, and immoral actions do not always have a common source, we would not respond by declaring that no action is any more unreasonable than any other, but by allowing that some unreasonable actions can be traced to breakdowns in something other than reason. There are errors in deduction and induction, ignorance of that of which one lacks experience, violations of simplicity and inelegance in the generation of hypotheses, failures of empathetic imagination, and—when a largely dissociable and so functionally distinct set of faculties are to blame—lack of resolve, deficits in prudential concern, and the absence of sufficient sympathy or humanity. Any of these failures can lead to a person's being properly termed "unreasonable" by the common speaker. Thus, were Hume to examine common usage with the requisite care, he would surely absolve the ordinary man of a criticism he primarily aims at the "greatest part of moral philosophy, antient and modern" (T 2.3.3.1; SBN 413). And the change in view would be decidedly superficial because Hume's primary concern was not the description of linguistic practice or the analysis of common concepts, but with eradicating the mistakes of those who assume that our normative talk reflects the mind's underlying structure.

Of course, this leaves open the question of whether Hume's detailed positive views of prudence and morality really do result in either nihilism or a wholly instrumentalist view of practical reason. And these questions surely demand more substantial answers. How, according to Hume, do we come to judge that a given action is weak-willed, imprudent or immoral? What leads us to conclude that we are obligated to refrain from such behavior or have good reasons to do so? And if these conclusions do not depend for their truth on the hypothesis that irresolution, imprudence and immorality result from impairments in reason, when are they truly or justly drawn? My aim in what remains of this essay is to explain why Hume's answers to these questions help clear him of Korsgaard's and Millgram's skeptical allegations. On the interpretation I will defend, Hume thinks that we are obligated to refrain from imprudence and immorality irrespective of what we might happen to want. And because Hume rejects "bare" or "ungrounded" obligations, he must allow that we have substantive non-instrumental reasons to act as we know we are obligated to act. We all have non-instrumental reasons to be prudent and moral.[31]

3. Hume's Account of Basic Prudence

We begin with prudence. A man uses his reason to figure out that his continuing to drink liquor to excess will likely end in a painful, overly precipitous death. Of course, this need not extinguish his appetite for the drink. When he looks at the beverage, and imagines the pleasurable glow that will result from its consumption—

not to mention the fun to follow—the attraction is significant. But he also has an aversion to harm and pain. When he imagines the future sufferings of liver disease and kidney failure, and is led by his understanding to think of them as the likely result of his drinking, he is filled with fear, and a consequent aversion to the actions that will bring the frightening scenario to fruition. If his aversion to the liquor is sufficiently strong, he will refrain from drinking. If it is not, and his craving is more intense, he will imprudently raise the glass to his lips.

Surely, cases of this kind are common enough, but are they all there is to prudence? Let us say that "Hobbesian" accounts of action deliver an affirmative answer to this question. Deliberation is nothing more than the alternation of attractions and aversions that result from imagining the pleasurable and painful aspects and consequences of our prospective actions. And the will is just the "final" appetite or aversion in this process, the "unseen beginning" of the action or forebearance that follows. If we equate the last appetite or aversion in deliberation with the one that is more intensely felt, we arrive at an account of prudence that assigns as little to reason as can be plausibly (or implausibly) maintained. Perhaps the prudent man more vividly imagines the future suffering that he knows will likely result from his drinking; or perhaps his understanding of the relevant causal or probabilistic relations better translates the image into an aversion to the drink before him; or perhaps the prudent man's lust for the booze is experientially weaker than the imprudent man's, either because the former tends to experience less pleasure from drinking, or because his memory does not convert his past pleasure into as intense a craving for the liquor he knows is its source. In any event, though the Hobbesian account does indeed distinguish the imprudent drinker from the man who prudently refrains, it does so on the basis of nothing more (nor less) than the relative phenomenological strengths of their respective passions. Reason's contribution to prudence is a genuine one, as it must be operating if the drinker is to represent his future suffering as the likely effect of his continued debauchery. But the understanding's role is limited to its effects on the balance of experienced aversion and appetite.[32]

Hume is quite clearly dissatisfied with the Hobbesian account. Of course he allows that an intense fear of suffering brought on by a vivid image of future pain can motivate us to resist a source of current pleasure.[33] But he also thinks that we can have a violent or phenomenologically robust craving for an immediate good that is motivationally weaker than a calm or phenomenologically anemic desire to avoid greater future pain.

> 'Tis evident passions influence not the will in proportion to their violence, or the disorder they occasion in the temper. . . . We must, therefore, distinguish betwixt a calm and a weak passion; betwixt a violent and a strong one. (T 2.3.4.1; SBN 418–9)

What then explains prudence? The Hobbesian account explains prudential motivation in terms of the felt strength of our aversion to a known future pain and its superiority in degree to the attraction felt to a present source of pleasure. Since Hume rejects the Hobbesean equation of motivational strength with phenomenological strength, he needs another explanation. Does he have one?

In fact, Hume expresses doubts as to whether we can fully explain prudence in more basic terms. Sometimes a man's calm passion for his long-term good prevails over a violent appetite for immediate pleasure; sometimes the reverse happens. All Hume initially says is that something about the person's "general character" or "present disposition" must be responsible. Indeed, Hume hypothesizes that the somewhat mysterious nature of prudence helps account for the inexact nature of psychological explanation and our common inability to predict each other's actions.

> Men often act knowingly against their interest: For which reason the view of the greatest possible good does not always influence them. Men often counter-act a violent passion in prosecution of their interests and designs; 'Tis not therefore the present uneasiness alone, which determines them. In general we may observe that both these principles operate on the will; and where they are contrary, that either prevails, according to the *general* character or *present* disposition of the person. . . . From these variations of temper proceeds the great difficulty of deciding concerning the actions and resolutions of men, where there is any contrariety of motives and passions. (T 2.3.3.10; SBN 418)[34]

We can say, in an Aristotelian vein, that training and habituation can aid us in our efforts to be prudent. A person who consistently forgoes present pleasure to avoid greater future pain will have concern for his overall good become "a settled principle of action," which as "the predominant inclination of the soul," will no longer produce any "sensible agitation" (T 2.3.4.1; SBN 418–9). This is, perhaps, one way to inculcate what Hume calls "strength of mind," which is Hume's label for the "prevalence of the calm passions over the violent" (T 2.3.3.10; SBN 418), though it will not augment strength of *will* as we customarily think of it. For on Hume's own description of the phenomena, habituation and training primarily work by diminishing the violence and force of imprudent desire, not by leaving that force in place and augmenting the opposing motivational strength of calm prudential concern.[35]

At any rate, Hume's initial discussion identifies two species of what we might call "basic prudence." A man will be prudent if he has a "general appetite to good, and aversion to evil consider'd merely as such," and either no violent passions for things that are bad for him, or extraordinarily weak passions inclining him to the lesser good (T 2.3.3.8; SBN 417). But a man can also be prudent if his concern for

his overall wellbeing is challenged by violent passions for things that are bad for him so long as he has the strength of mind necessary to resist their "solicitations" (T 2.3.3.10; SBN 418). Perhaps the second form of prudence is somewhat less basic than the first, but Hume's initial account gives us no reason to suspect that either need involve much in the way of reflection or reasoning. Perhaps, as on the Hobbesian account, we only need the faculty of reason to be prudent because without it we would be unable to determine which actions will yield our overall happiness, pleasure, or good. Perhaps our concern for that good and the strength of mind we need to overcome internal obstacles to its attainment, have their source in something entirely distinct from reason.

4. The Inadequacies of the Basic Account

Were we to end our description of Hume's account of prudence here, at T 2.3.3, we would be forced to join Korsgaard and Millgram in a substantially skeptical reading. All the understanding does, according to what is perhaps the most common interpretation of Hume, is help us identify means to our ends; reason can do nothing to aid us in our efforts to settle upon worthy goals. While attributing this minimal view of practical *reasoning* to Hume is consistent with granting him a more robust view of *reasons* for action, it does place that project in an unfavorable light. For suppose, through no fault in his reason, the drunkard loses all concern for his future wellbeing: How can we say that he has (non-instrumental) reasons to maintain his health, if the only thing his reason can "tell" him is how drinking will frustrate the satisfaction of ends he just does not have?

But the received reading is exceedingly uncharitable, and the account we have sketched to this point is importantly incomplete. For one thing, it entirely omits Hume's views on how we come to *evaluate* basic imprudence. Of course, as we have seen, Hume insists that it need not be "unreasonable" to lack concern for one's future wellbeing, nor unreasonable to lack the strength of mind necessary to secure one's greater good (T 2.3.3.6; SBN 415–6). But he also insists that prudential concern and strength of mind are *virtues*, and their absence a *vice*. And since imprudence is a vice, Hume thinks it appropriate to say that we *ought not* be imprudent, indeed, that we are *obligated* to do what is good for us.

> When any action, or quality of the mind pleases us *after a certain manner*, we say it is virtuous; and when that neglect, or non-performance of it, displeases us *after a like manner*, we say that we lie under an obligation to perform it. (T 3.2.5.4; SBN 517)

Importantly, though Hume argues that the viciousness of the vices lies in the "breasts" of those who survey them with the requisite disinterested sympathy,

there is no point at which Hume says that prudence and weakness of mind are not vices, or that they are only vices in those whose ends they frustrate, much less that we have no reason at all to be prudent and resolute. If Hume thinks that we ought not drink to excess, and that the unrepentant drunk ought to put down his drink, indeed if he thinks it is appropriate to say that we are all obligated to seek the balance of pleasure over pain, isn't Hume committed to saying that we all have reason to practice moderation? Korsgaard and Millgram must fully integrate Hume's account of normative judgment with his account of moral and prudential action if they are to adequately justify their skeptical readings.

Secondly, Hume is quite clear that we can evaluate our own characters and actions in the same terms we use to evaluate the characters and actions of others.[36] What, according to Hume, will happen when the drunk reflects on his own intemperate passions and forms a judgment as to their character? Mightn't he explicitly judge that he ought not take the drink because doing so would be imprudent? And mightn't he come to think that the imprudence of his prospective action is a good reason not to perform it, even if, up until the very moment of this realization, he lacked an adequate concern for his future good? If so, what role might his judgments play in reinforcing his calm passion so that he might act prudently? After all, though Hume insists that "[g]enerally speaking, the violent passions have a more powerful influence on the will," soon after T 2.3.3 he concedes that, "'tis often found, that the calm ones, when corroborated by reflection, and seconded by resolution, are able to controul them in their most furious movements" (T 2.3.8.13; SBN 437–8). Unless we come to terms with Hume's account of the relevant sort of *reflection* and *resolution* we cannot pretend to have offered an adequate interpretation of his views on practical reason.

5. Hume's Account of Evaluation[37]

Hume's theory of evaluation begins with an explanation of the criticism we direct toward others. Suppose, then, that we observe the drunkard ingest fifteen pints of strong beer in a half-hour period. While this feat might arise from a blameable lack of concern for his future wellbeing, our man might be competing in a drinking contest that he knows he must win to keep his family from starving. Somewhat less fancifully, without knowledge of the man's past, we must entertain the possibility that this is a one-time binge, an exception to his general temperance. Hume rightly insists that whether we resent or blame an action depends on the motives to which we assign it. And, when we do condemn an action, the kind of sentiment we experience is similarly sensitive to motive.[38] In all events, Hume is clear that when we attribute excessive consumption to the vice of imprudence we will have used our understanding to establish it as the most likely cause of the actions under review.

> 'Tis evident, that when we praise any actions, we regard only the motives that produc'd them, and consider the actions as signs or indications of certain principles of mind and temper. The external performance has no merit. We must look within to find the moral quality. This we cannot do directly; and therefore fix our attention on actions, as on external signs. But these actions are still consider'd as signs; and the ultimate object of our praise and approbation is the motive, that produc'd them. (T 3.2.1.2; SBN 477)[39]

So suppose that we have verified that the drunkard is drinking to excess because, though he has made an accurate assessment of the damage he is inflicting on himself, he lacks a sufficiently robust "appetite for his good as such." Will this on its own allow us to conclude that his actions are vicious? Famously, Hume insists otherwise.

> Take any action allowed to be vicious: Willful murder, for instance. Examine it in all lights, and see if you can find that matter of fact, or real existence, which you call vice. In which-ever way you take it, you find only certain passions, motives, volitions, and thoughts. There is no other matter of fact in the case. The vice entirely escapes you, as long as you consider the object. You never can find it, till you turn your reflection into your own breast, and find a sentiment of disapprobation, which arises in you, towards this action. Here is a matter of fact; but 'tis the object of feeling, not of reason. It lies in yourself, not in the object. So that when you pronounce any action or character to be vicious, you mean nothing, but that from the constitution of your nature you have a feeling or sentiment of blame from the contemplation of it. (T 3.1.1.26; SBN 468–9)

We must *feel* blame toward imprudence if we are to classify it as a vice. Still, one might ask, what makes us feel a pleasurable sort of love or admiration when we consider prudence, and a negative sort of hatred or disapproval when we consider imprudence? Why don't our sentiments align themselves in the opposing directions? More generally, why don't we feel moral admiration toward those things we call "vices," and disapproval towards those we call "virtues"?

To answer these questions Hume again invokes the understanding. When we investigate the matter more fully we find that moral disapproval tracks characteristics that are in general harmful and pernicious, and moral approval tracks the opposite sort of trait. In fact, my sentiments of moral approbation track something other than just my own wellbeing, as they are significantly distinct from the approval I feel for things I judge to be in my self-interest. We have our own pleasure

or happiness in mind when we judge that someone's character will likely propel them to act in a way that is *good for us*, but we have either that person's own interests in view or the interests of those in his immediate circle, when we judge the characteristic a *virtue*.

> Every quality of the mind is denominated *virtuous*, which gives pleasure by the mere survey; as every quality, which produces pain, is call'd *vicious*. This pleasure and pain may arise from four different sources. For we reap pleasure from the view of character, which is naturally fitted to be useful to others, or to the person himself, or which is agreeable to others, or to the person himself. One may, perhaps, be supriz'd, that amidst all these interests and pleasures, we shou'd forget our own, which touch us so nearly on every other occasion. But we shall easily satisfy ourselves on this head, when we consider, that every particular person's pleasure and interest being different, 'tis impossible men cou'd ever agree in their sentiments and judgments, unless they chose some common point of view, from which they might survey their object, and which might cause it to appear the same to all of them.[40]

Thus, if we are to judge that the drunk is being imprudent, and that his imprudence is a vice, we must put reason to work twice over. First we must infer the man's motive from his actions. Then, after we have figured out that he is in part moved by a general lack of concern for his future, we must rely on our background knowledge that this trait is not naturally fitted to be "useful to others, or to the person himself," nor "agreeable to others or the person himself," where this knowledge will almost always owe its existence to the prior operations of the understanding.[41]

Why then must sentiment enter at all? Isn't our knowledge that a characteristic such as imprudence is both destructive to a man and his circle and disagreeable to their survey enough to license us in judging it to be a vice? And if these considerations do justify that judgment, why cannot reason alone inform us that we should be prudent because imprudence is a vice?

Well, at least in paradigm cases of moral judgment, Hume insists that it is not enough to know that the drunk's behavior will cause both his ruin and the ruin of those who depend on him for support. Particularly in the *Treatise*, Hume emphasizes that we must be endowed with what he calls "sympathy" if our sentiments are to be affected by the suffering we subsequently imagine the drinker and his dependents will experience. Indeed, so called "artificial virtues" such as a sense of justice—conceived of as respect for laws enshrining rights to property—need not benefit those who have them and need not benefit those in their immediate circle (T 3.3.6.1; SBN 618). Artificial virtues are therefore only denominated virtues because they benefit society on the whole and because we sympathize (to at least

a minimal extent) with our compatriots.[42] But sympathy also plays a role in our approval of natural virtues like prudence.

> Most people will readily allow, that the useful qualities of the mind are virtuous, because of their utility. . . . Now this being once admitted, the force of sympathy must necessarily be acknowledg'd. Virtue is considered as a means to an end. Means to an end are only valu'd so far as the end is valu'd. But the happiness of strangers affects us by sympathy alone. To that principle, therefore, we must ascribe the sentiments of approbation, which arises from the survey of all those virtues, that are useful to society, or to the person possess'd of them. These form the considerable part of morality. (T 3.3.6.2; SBN 618–9)

It is only if we are endowed with a sensitivity to the sentiments and welfare of others that we will be able to move from our understanding that a character is both unseemly and lacks utility to the impression of hatred or disapprobation that constitutes our moral disapproval of it. And it is only if we feel such hatred or disapprobation in our breasts, that we can formulate distinctively moral judgments.[43]

6. Reflection and Resolution

With Hume's view of moral criticism on the table we can now reconsider our two varieties of basic imprudence. Suppose, again, that we have used reason to determine that our man drinks excessively because he either lacks a sufficiently robust concern for his future, or, though he is genuinely torn between the alternatives, he lacks the resolve necessary to overcome his violent passion for liquor. These traits are, perhaps, immediately disagreeable. More importantly, if we properly grasp the relevant causal facts, and stop to think of the matter, we will find ourselves contemplating the pain and suffering that both the drunk and his loved ones will likely suffer from his continued inebriation. Since, as normal human beings, we are endowed with some degree of sympathy, imagining their pain will make us feel bad; and then, through a complex process involving both sympathy and imagination, the negative affect we experience will transform itself into a sentiment of blame aimed at the man responsible for all this suffering. Once this sentiment is experienced, we will judge the man's lack of concern or resolution a vice, and (insofar as we are inclined to express our views on the matter) tell him he should change his ways.

Now, as far as I can tell, there is no reason to think that this explanation of how we come to critically evaluate imprudent people imperils the truth or propriety of our criticism. Of course, on Hume's account, the judgment that people should not act imprudently is not the product of reason alone. But it cannot be reasonably

maintained that a judgment must be wholly a priori to be true or warranted. And though Hume has our impressions of reflection playing an essential role in critical judgment that they need not play in ordinary causal inference, these sentiments also play an essential role in our wholly non-evaluative beliefs about them; and the function of sentiments in moral judgment is not wholly unlike the function of impressions of sensation in the generation of causal belief. Nor is there anything in Hume's explanation of our critical judgments that should lead us to think that we can only legitimately condemn others when they fail to take means they know (or reasonably believe) they must to achieve their ends. So the skeptical and instrumentalist readings already have an air of implausibility to them. And this suspicion is only reinforced when we see Hume adopt a dismissive attitude toward moral skepticism, and describe his theory, on which morals are grounded in sympathy, as "pleasing" to "all lovers of virtue."[44] It is true that Hume's account of virtue and vice is significantly "anti-realist" in its requirement that the sentiments or emotional reactions of an observer must be added to a harmful trait if that trait is to really constitute a vice. But the traits and disutilities under review are real enough.

In fact, on Hume's theory, reason or the understanding commonly *corrects* the moral sentiments so that they better track the utilities and disutilities in question. And because Hume thinks this process of correction is fallible, he allows that our moral judgments can be *inaccurate*. I might, for instance, fail to properly adopt the perspective of a trait's bearer or his circle and so judge someone vicious merely because he frustrates my individual self-interest. In such a case, Hume says, my blame is misplaced, and I have failed to "give praise to what deserves it" (T 3.1.2.4; SBN 472).[45] In a different sort of scenario, I might infer that a person's character really is vicious because it is harmful to both its bearer and his circle, but then fail to feel blame or disapproval toward him because my sentiments never properly "adapt" to my realization. For example, being infatuated with a woman, I might be incapable of resenting her in any experientially salient way. But this needn't inhibit me from using my reason to determine that I *would* feel this way were I not in love, as her character and actions are uniformly rotten. Indeed, this is precisely the sort of counterfactual reasoning Hume invokes to explain how we maintain consistency in our moral judgments across temporal and spatial distances.

> We blame equally a bad action, which we read of in history, with one perform'd in our neighbourhood the other day: The meaning of which is, that we know from reflection, that the former action wou'd excite as strong sentiments of disapprobation as the latter, were it plac'd in the same position. (T 3.3.1.18; SBN 584)

It seems that if I am sufficiently smitten, but nevertheless capable of the kind of "reflection" Hume mentions here, I might appropriately judge that my beloved

is bad, and be entirely justified in asserting that she is bad, even though I cannot bring myself to experience the negative emotions or sentiments one might infer from my speech. As Hume says,

> The passions do not always follow our corrections; but these corrections serve sufficiently to regulate our abstract notions; and are alone regarded when we pronounce in general concerning the degrees of vice and virtue. (T 3.3.1.21; SBN 585)[46]

Admittedly, it is hard to find passages in which Hume explicitly says that we can classify people and their motives and actions as vicious or virtuous without experiencing the sentiments through which we draw moral distinctions in the first instance. But his positive theory of moral criticism surely predicts that this will happen. After all, on Hume's account of it, sympathy is an imperfect and somewhat irregular principle; and though it works in concert with the understanding, the two faculties are wholly distinct in nature. So it should be possible for someone with "acquired psychopathy" to entirely lose his sympathy while retaining both his memory of its operations and the understanding he needs to continue discerning which traits and actions are harmful.[47] We might suppose that such a man knows that his actions are harming others, and knows that were he still endowed with sympathy he would be reacting to his victims' suffering with negative sentiments. We might even suppose that he remembers how hurting others led him to experience a form of shame and self-censure in the past, and how these reactions would then lead him to think of his own actions as vicious. Could he not then come to judge that he must be acting in a vicious way now, since there are no relevant differences between the two cases? And mightn't he do so without grounding his knowledge of his own viciousness in a current feeling of remorse or disapprobation? Hume never straightforwardly addresses the possibility, even when discussing entire communities, like the Scythians, who lost their sympathy or humanity from an over-emphasis on martial bravery and courage.

> The Scythians, according to Herodotus, after scalping their enemies, dressed the skin like leather, and used it as a towel; and whoever had the most of these towels was most esteemed among them. So much had martial bravery, in that nation, as well as in many others, destroyed the sentiments of humanity; a virtue surely more useful and engaging. (EPM 7.14; SBN 255)

Nor does Hume explicitly assert the possibility of moral judgment in the absence of moral motivation when describing individuals like Timon "the manhater" who was reportedly, "worse than indifferent to the images of virtue and vice," instructing

Alcibiades, "Acquire the confidence of the people: You will one day, I forsee, be the cause of great calamities to them" (quoting Plutarch at EPM 5.2.25; SBN 226–7). Of course, the social phenomenon exemplified by the Scythians is importantly different from that in which a single individual lacks sympathy. For when the loss of humanity occurs throughout an entire community, it seems unlikely that its members will be able to retain a proper set of moral classifications. But it remains that Hume never says whether, despite having sentiments inverted in relation to those of his compatriots, Timon knew right from wrong.

Still, Hume may have been thinking of Timon as a *congenital* psychopath—someone who *never* experienced those sentiments without which (on Hume's view) moral distinctions cannot be drawn. So let us suppose, to put a contemporary name to a decidedly late-blooming psychopath, that it is only after Phineas Gage has a pole run through his prefrontal cortex that he is incapable of feeling anything at all when he perceives or imagines the suffering of others.[48] Now it might be suggested that on Hume's account Gage's loss of sympathy or humanity would of course rob Gage of his ability to formulate genuinely moral judgments. Before his accident Gage used "Cruelty is vicious" to give voice to his *impassioned* belief in the viciousness of cruelty, but if Gage now lacks all fellow feeling, he must use "Cruelty is vicious" in its "inverted commas sense"—as R. M. Hare would describe it—to give voice to his cold knowledge that harming others without concern for their wellbeing is a *so-called* vicious way to behave, or that cruel behavior is the kind of thing called "vicious" by the typical member of his community.[49]

But though Hume might have joined Hare in this view, his subscription to the doctrine is nowhere evident in the text. For although Hume argues that in the first instance vice is more properly "felt" than judged (T 3.1.2.1–3; SBN 470–1), we have already seen how, on Hume's account, counterfactual reasoning can loosen the capacity for moral judgment from its roots in our emotions. Reconsider, for instance, the case above, in which I judge that the object of my infatuation is vicious even though I am entirely unable to feel anything remotely censorious toward her. On Hume's account, I needn't relegate myself to judging that the woman is called "vicious" or that she is so-called vicious. Instead, counterfactual reasoning allows me to *use* rather than *mention* "vicious" and in so doing judge that the woman I love really is vicious. Similarly, on Hume's reckoning, Gage could have no idea of the viciousness of cruelty if he never experienced hatred for the vice when surveying it from the general point of view. But Gage needn't rehearse this experience (or instance the requisite impression) on every occasion on which he thinks of an action as vicious. So while Gage might respond to his deficit by limiting his judgments to linguistic matters, he needn't. Up until his accident he thought of intentionally harming others for self-interested reasons as vicious, and he can recognize that the rather dramatic brain damage he has suffered is the only relevant difference between now and then. So he will likely infer that cruelty must

still be vicious, and that his inability to feel the sentiments that (on Hume's view) make cruelty vicious is the product of horrible damage to his brain—damage that has in turn robbed him of the means by which he first distinguished the virtues from the vices. More importantly, it is far from obvious that Hume's view of the mechanisms responsible for a healthy man's moral judgments would have led him to conclude that Gage could not reason in this way. (Indeed, his discussion of the role played by counterfactual belief in normal moral thinking suggests the opposite verdict.) And if Gage did reason in this way, it is hard to see why Hume would feel compelled to accuse the man of mentioning rather than using "vicious."[50]

At any rate, our pitiful drunkard needn't be this far gone. He might find himself with a calm passion for his greater good and a violent passion for the drink before him, and turn to reflection and resolution to see him through. And, on Hume's account of them, his reason will play a role in his efforts comparable to the one it plays when we survey him from afar. He must think not just of the effects of his drinking on his future self, but consider the other people his actions will harm. He must consider his want of resolution and lack of concern for the future as general traits of character, and gauge the multifaceted impact they will have on his life and the lives of those he regularly encounters. And he must imagine his future existence and the existence of others from the "inside" so that his sympathy might be engaged by the suffering he is likely to cause if he persists on his current course. Indeed, so long as our man is adequately endowed with understanding and humanity, Hume is willing to allow that these reflections can tip the scales in favor of his calm passion and play a role in causing him to act prudently.

> But may not the sense of morality or duty produce an action, without any other motive? I answer it may. . . . When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle, may hate himself upon that account, and may perform the action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least to disguise to himself as much as possible his want of it. (T 3.2.1.8; SBN 479)

So the kind of strength of mind that distinguishes the prudent man from the imprudent is not entirely mysterious. In the end, Hume does not just throw up his hands and declare that sometimes our calm passions win out, whereas sometimes the current temper holds sway, we know not why. There are at least certain cases in which self-evaluation and self-censure play a causal role in moving us to act prudently or morally in the face of temptations to the contrary. And when we examine Hume's account of the relevant kind of evaluation and censure, we find reason operating in spades, albeit reason augmented by sympathy or humanity.

What Hume would not accept, I think, are Kant's rationalist conclusions regarding the power of the good will. Since we judge that a man like Gage, who feels nothing for others, nevertheless *ought* not harm those who stand in the way of his often selfish ends, Kant thinks we must conclude that Gage *can* refrain from immoral behavior. Indeed, the psychological sense of possibility that figures in Kant's reasoning is really quite strong: Gage must be able to bring himself to do the right thing without having a gun put to his head or being screamed at or bullied. But given the destruction of his sympathetic faculty, Gage cannot be made to feel bad for the pain and suffering he is causing. We are therefore in a position to conclude that the man must be able to act from reason with no help from either external coercion or his emotions or sentiments. By merely reflecting on the fact that he is acting on a maxim or principle that cannot be coherently universalized, or by reflecting on the fact that he is failing to treat the humanity of others as something more than just a means to his ends, he must be capable of acquiring a motive strong enough to overcome his drive to realize his goals at all costs. His reason alone must have the power to generate a certain kind of motivationally efficacious respect or esteem for the moral law.

> Thus suppose the mind of [a] friend of humanity were clouded over with his own grief, extinguishing all sympathetic participation in the fate of others; he still has the resources to be beneficent to those suffering distress, but the distress of others does not touch him because he is sufficiently busy with his own; and now, where no inclination any longer stimulates him to it, he tears himself out of his deadly insensibility and does the action without any inclination, solely from duty. (Ak 4:398; *Groundwork*, 14)

Reason alone must be capable of generating a motive, and therein subverting immoral passions that would otherwise hold sway. Reason must be master, not slave.[51]

7. Practical Reasons

We are now in a position to extract a view on reasons for action from Hume's accounts of prudence and morality. On the traditional reading, Hume thinks that we only have reasons to be prudent or moral when prudential and moral behavior provides an effective means to ends established by our passions. Should we endorse the traditional reading? Should we even go so far as to agree with Derek Parfit's claim that Hume denies the very the existence of practical reasons (128)?

The issue is complicated by the incredible diversity in our use of "reasons." First, "reasons" can just mean *causes* as when one sets out to state the reasons why the Earth rotates on its axis without trying to psychologize or justify the planet's

behavior. Second, "reasons" can be used to pick out what philosophers often call "normative reasons"—i.e., *pros* and *cons*—as when an inquirer asks if there are any reasons to think the Mets will win the World Series in a context in which it is clear that he as yet has no beliefs on the matter. Finally, "reasons" can be used to denote *psychological states* that explain a person's actions or attitudes. If asked, "What were his reasons for breaking into our empty garage?" one can properly answer, "Well he needed cash to feed his habit, and thought we owned a fancy car," therein citing a desire and belief to "rationalize" an action done for no good (normative) reason. Now I take it that it is the second of these three uses of "reason" that is in play when we ask whether Hume denies the existence of non-instrumental reasons for action. The question is not whether Hume believed in beliefs, desires, and causal relations. The question is whether we can truly or legitimately cite considerations in favor of someone's acting prudently and morally that go beyond facts about which actions he must perform to realize ends antecedently put in place by his appetites, aversions, and passions.[52]

Now if we ask the question in this way, I think Hume's answer to it would, indeed, posit non-instrumental reasons to be prudent and moral. After all, when we tell a man that he shouldn't continue to drink to excess because doing so would be imprudent, we give him a reason to refrain from drinking; and, on Hume's view of the matter, we can do this in a way that goes beyond our merely pointing out means to his antecedently given ends. If the drunk comes to recognize that his actions are imprudent by actively considering their real or likely effects, and (in Kant's words) "sympathetically participating" in his imagined future suffering or the suffering he knows his dependents must experience, he can come to acquire an end that he has not yet fully set for himself. That is, before hearing our counsel he was not actively engaged in pursuing his greater good; afterwards he might very well be. And we will have gotten him to pursue the prudent course not by telling him of the effects of excessive drinking, effects, we might assume, of which he was already in some sense aware. Instead, he will have been led to embrace his long-term wellbeing by vividly dwelling on the viciousness of intemperance. Surely, telling a man who is already sufficiently concerned for his health that drinking causes illness is one thing; getting a drunk who knows about the effects of alcohol to care more about his health is another.

Of course, suggesting to a man that the viciousness of intemperance is a reason for him to stop drinking would be fairly pointless if there were no prospect of his being moved by that consideration to take steps to mitigate his consumption of booze. And, if Hume is right, when we assume that moral considerations can move others to change course, we are relying on their having something more than an adequately functioning faculty of reason. What, then, should we say about agents, like Timon the manhater, who (we might suppose) have *always* lacked the sympathy and humanity Hume thinks a person needs if he is to move from an abstract

appreciation of harm and disutility to an emotionally engaged perspective from which moral action might result? In order to say that the congenital psychopath has distinctively moral reasons to change his ways, must we join Kant in attributing motivational powers to unaugmented reason? Though the issue is difficult to resolve, I think that in answering this question Hume would indeed have to come to adopt a somewhat "externalistic" perspective on reasons for action.[53]

To illustrate what I have in mind, I want to consider a future in which a new surgery becomes available whereby color vision can be instilled in the congenitally colorblind. And I want to consider how, as a color-sighted citizen of this future, you might try to convince a man who has never seen a color that he should sign up for the operation. Surely, in trying to steer him toward the surgery, it would be best not to use color concepts. You might do this, for instance, by telling him that after the operation he will be better able to discriminate in ways that matter to the vast majority of us who have color vision; he will be better able to dress, decorate, and paint so as to please both his newly color-sighted self and the other color-sighted people he will encounter throughout his days. If you were to instead flatly insist that the red of a sunset and the blue-green of the sea are worth seeing, he would not be able to adequately grasp what you had said, much less appreciate the considerations you had advanced as reasons to acquire color vision.[54] Does this mean that the fact that red sunsets are beautiful is not a reason for this man to acquire color vision? Must we, as color-sighted people, adopt the perspective of the colorblind when determining what reasons they have to see colors? Or can we justifiably or even truly say that there is an entire class of reasons to view the colors that just could not play a role in affecting the decisions of a fully colorblind person?

I think Hume's positive doctrines strongly suggest a positive answer to this last question, and that the tenor of his discussion of moral skepticism is evidence that he would have given a similar answer to whether creatures with understanding but neither sympathy nor humanity have reasons to be moral. If we want to effectively communicate with someone who has never felt anything for others, we must give him self-interested reasons to be moral. We must speak of the punishments of the jailhouse or the fires of hell—invoke the rewards of a good reputation or the luxury of eternal paradise. But these are not the main reasons to respect other people, much less the only reasons to do so. Of course, if Hume is right, when we tell the congenital psychopath that he is being cruel or unkind, he will not adequately grasp what we have said, and so will not understand that the cruelty of his behavior is a reason to stop engaging in it. But we need not adopt the *psychopath's* perspective when laying out the reasons why he should not harm others. When *we* think about why he should not act immorally, we can legitimately use the very concepts of virtue and vice that inform all coherent moral thought and discourse. We simply cannot legitimately expect to share our thoughts with him in a way that might affect his behavior.[55]

Why do I think that Hume would endorse this way of thinking about practical reasons? Well, when he does explicitly talk of reasons for action, he does not treat them as purely instrumental. Instead, he seems to assume that the fact that an action would be pleasurable is a non-instrumental reason to pursue it, and that the same goes for the fact that an action is demanded by virtue. In truth, Hume talks as though pleasure, pain, virtue, and vice provide *foundational* reasons for action—reasons to act for which a normal human being needs no reason.

> Ask a man why he uses exercise; he will answer, because he desires to keep his health. If you then enquire, why he desires health . . . he may . . . reply, that it is necessary for the exercise of his calling. If you ask, why he is anxious on that head, he will answer, because he desires to get money. If you demand Why? It is the instrument of pleasure, says he. And beyond this it is an absurdity to ask for a reason. . . . Something must be desirable on its own account . . . virtue is an end and is desirable on its own account, without fee and reward, merely for the immediate satisfaction. (EPM App.1.18–20; SBN 293–4)

Good or pleasure is desirable "on its own account" and so provides us with a foundational reason to pursue money—a reason for which we need no reason, a reason that remains even if we set our sights on acquiring pain rather than pleasure because we prefer a "lesser good" to one that we know to be greater. Moral reasons have a similarly non-instrumental source. I ought to save a child drowning in shallow water because inaction would be mean and meanness is a vice. And this, Hume would say, is an appropriate place to stop giving and asking for reasons. The virtues and vices are "regress stoppers."

Of course, even though, on Hume's view, a normal human being will not need a reason to refrain from cruelty, we, as theorists, can still ask wholly "descriptive" questions about what leads us to think of cruelty as a vice. But, as should now be clear, Hume's answers to these questions do not portray us as inferring the viciousness of meanness from its role in thwarting a mean man's ends. Instead, meanness is considered a vice because it invokes disapproval from us when we adopt the requisite perspective or set of perspectives. And it will remain true that a particular agent's meanness invokes this disapprobation even if, because that agent is vicious, he has no desires that will be served by his rescuing the helpless child we have described. Indeed, we can properly say that a mean man has a reason to intervene even when he refuses to occupy the perspectives of those affected by his meanness, and so fails to acknowledge his viciousness out of culpable ignorance. And we can surely say something similar when he admits that failing to intervene would be mean, but out of indolence or cowardice refuses or fails to do what he knows he has good reason to do. Thus, while contemporary "Humeans" tend to

allow instrumental norms and reasons, but no reasons that are not derived from them, a charitable interpretation of Hume would have him acknowledging non-derivative norms and reasons of self-interest, prudence, and morality.[56]

Why then is Hume so often read as either an instrumentalist or skeptic? I have suggested that one source of confusion may be a misplaced emphasis on T 2.3.3 and a consequent failure to augment the practical uses of reason that Hume describes there with its role in generating self-evaluation and strength of mind. But perhaps another explanation can be found in Hume's famous discussion of the "sensible knave" who thinks that in certain cases "an act of iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy" (EPM 9.2.9; SBN 282–3). Does Hume's response to the knave betray an instrumentalist conception of normativity?

The first thing to note is that the passage in question occurs in the midst of a section of EPM dedicated to the relation between morality and self-interest. That is, at the outset of EPM 9.2, Hume informs us that he is leaving consideration of the distinctively moral value of our actions behind to consider what wholly self-interested reasons we might have to be moral.[57] And it is entirely consistent for Hume to say that though the knave has no self-interested or instrumental reasons to respect his contracts and the property of others, he nevertheless ought to do so because it is vicious not to, where the viciousness of his plans supplies him with a reason against their execution. The second thing to note is that the knave under consideration is not presumed to be cruel or hateful, but just wily and deceitful. These are, on Hume's reckoning, "artificial" vices, which derive their immorality from conditions peculiar to civil society. Given Hume's distinction between the natural and artificial vices, it would be entirely consistent for him to say that we all have non-instrumental reasons to refrain from cruelty and maliciousness, but that only those who have a sufficient concern for the good of society as a whole have a reason not to lie and steal.[58]

But, in fact, Hume does not say this. Instead, he acknowledges that it may be in the knave's narrow self-interest to cheat when he can get away with it,[59] but that the actions in question are nevertheless villainous and base. And the claim that follows this description comments on the absence of a moral *motive* in this case, not a moral *reason*.

> I must confess that, if a man think that this reasoning much requires an answer, it will be a little difficult to find any which *to him* appear satisfactory and convincing. If his heart rebel not against such pernicious maxims, if he feel no reluctance to the thoughts of villainy or baseness, he has indeed lost a considerable *motive* to virtue; and we may expect that his *practice* will be answerable to his speculation. But in all ingenuous natures, the antipathy to treachery and roguery is too strong to be

> counterbalanced by any views of profit or pecuniary advantage. (EPM 9.2.10; SBN 283; emphasis added)

Hume argues that we should not cheat because such behavior is pernicious in a civil society where respect for property is both necessary and good over the long term. And he seems comfortable with claiming that we can rationally, appropriately, even truly persist in saying this to a knave we know lacks the humanity or fellow feeling he would need to be moved to change his ways through exposure to our criticism. Thus, Hume's discussion of the knave seems to have him adopting in practice the somewhat "externalist" perspective on reasons for action that I have said he would have embraced were he to have given an explicit account of such reasons. The man shouldn't lie and cheat; but if his heart is hardened, and he is sufficiently crafty, he probably will.

Instrumentalists will not be satisfied with this stance. But why not? I think that their dissatisfaction may lie in a common (though I think ultimately mistaken) view according to which there are three distinct senses of "ought" codifying three different sorts of normative authority.[60] That is, many theorists think that "ought" is used to express *self-interest* when we say that a man ought to perform actions he knows he must if he is to accomplish goals he is unwilling to abandon; that "ought" is used to express the distinct concept of *prudence* when we say that a man ought not do those things that will cause him harm; and that "ought" is used in the still distinct sense of *moral obligation* when we say that everyone ought to refrain from cruelty and injustice. Importantly, on the standard account, the principle enshrining our duty to adopt known means to our ends is *analytic* and so known to be binding via reason alone operating in a fairly straightforward way.[61] Whereas knowledge of moral and (perhaps) prudential "ought"s is problematic from the instrumentalist's perspective (and would have to be granted the status of synthetic a priori), we need do nothing more than reflect on the concepts, definitions, or "relation of ideas" that the instrumental principle brings to mind to know that it is both true and authoritative. Thus, to show that moral and prudential norms really are true or valid, we must somehow derive them from the instrumental norm and a posteriori knowable facts about the conditions of civil society. To show that a man *really* ought to be prudent and moral—or that he always has a good *reason* to act in these ways—we must show that imprudent and immoral behavior must always run contrary to a man's considered ends.[62]

Now I think that Korsgaard and Millgram quite accurately perceive that this is not Hume's view of the matter. On Hume's account of it, the claim that we ought to adopt known means to our ends is nothing more than a discursive representation of the viciousness of weakness of mind. And the viciousness of this characteristic—like the viciousness of imprudence, cruelty, and injustice—is felt and then judged to obtain a posteriori through the exercise of reason augmented by humanity. (We

must imagine what it is like to lack resolution, and what it is like for those who depend for their happiness on the irresolute, and then react in an appropriately valenced way to the negative emotions we are sure to imagine.) So on Hume's view there is no normative principle that can be rightly attributed to reason alone: the purported "ought" of self-interest does not have this status, nor does the "ought" of prudence, nor the "ought"s and "ought not"s that might be used to capture the other commonly recognized virtues and vices. But Korsgaard and Millgram are wrong to represent Hume's thoroughgoing rejection of principles of *pure* practical reason as tantamount to skepticism. Instead, Hume thinks that the good practical norms we have mentioned are all valid principles of *impure* practical reason: that is, the understanding operating in concert with sympathy. And though there may be problems with his effort to steer a course between rationalism on the one shore and instrumentalism and nihilism on the other, the project is not wrongheaded in any obvious way. Surely we need not think of reasons to act in such a way that we define Hume's attempt into incoherence.

NOTES

I would like to thank Francis Dauer, Alan Hájek, Elijah Millgram, Michael Rescorla, Sydney Shoemaker, and audiences at Claremont Graduate University and the Research School of Social Sciences at the Australian National University for helpful discussion. (Terence Irwin and Nicholas Sturgeon have also exerted a profound influence on my understanding of the texts.) I owe special thanks to Don Garrett, Thomas Holden, and several anonymous referees for *Hume Studies* for extremely useful written comments on earlier versions of this essy.

1 Throughout this essay I cite Hume's texts using the book, part, section, and paragraph number format made possible by some fairly recent editions of his work: *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2000); *An Enquiry Concerning Human Understanding*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 2000); and *An Enquiry Concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 2000). The title of the first of these enquiries will be abbreviated "EHU," the title of the second "EPM," and the title of the *Treatise* "T." I have also included references to the traditional editions of the *Treatise* and *Enquiries* prepared by L. A. Selby-Bigge and revised by P. H. Nidditch. These are denoted by "SBN" and followed by the relevant page number or numbers.

2 Does not passion need reason as much as reason needs passion? The Davidsonian (a sort of neo-Humean) insists that beliefs must be added to desires to produce action, so that reason and passion are coequal; neither is enslaved to the other. But when interpreted as an exceptionless generalization, the Davidsonian dictum is implausible. There are so called "basic actions" for which means/ends beliefs cannot be found, and when these actions are mere expressions of passion and quite obviously fail to achieve any of the agent's ends, he need not believe that they are in any way desirable. (Striking a wall

in anger is a good example of this phenomenon.) Still, it is not entirely clear whether Hume views purely passionate action as a possibility. Though he posits productive passions that do not owe their existence to reasoning at T 2.3.3.8 (SBN 417) and 2.3.9.8 (SBN 439), he never says that these passions can yield action without the aid of reason. Indeed, he seems to deny it when he says, "Human nature [is] compos'd of two different parts, which are requisite in all its actions, the affections and understanding" (T 3.2.2.14; SBN 492). Suppose then that Hume thinks that the understanding is in fact necessary for action. I will argue below that Hume also assigns the understanding an essential role in our evaluation of our own actions and characters as virtuous or vicious, prudent or imprudent—where this kind of evaluation regularly results in a suitable modification of our plans and behavior. If this reading is correct, then the enslavement metaphor is at least overblown. Reason is not limited to discovering the means to our ends; though it cannot set ends on its own, it plays an important role in the self-conscious, reflective adoption of prudent or moral goals. Some commentators have gone so far as to deny that Hume thinks passions need play a role in every action; see, e.g., Ingmar Persson, "Hume—Not a 'Humean' About Motivation," *History of Philosophy Quarterly* 14.2 (1997): 189–206, and Nicholas Sturgeon, "Moral Skepticism and Moral Naturalism in Hume's Treatise," *Hume Studies* 27.1 (2001): 3–83. The best evidence for the Persson/Sturgeon case comes at T 1.3.10 (SBN 118–23), where Hume says that "not every idea" of pain or pleasure has an equal effect on our actions, and clearly implies that some ideas do have an effect. But I read Hume as here distinguishing believing that some action will cause pleasure from merely imagining that it will, and see him as consistently holding that (of this pair) the belief alone will lead to action but only when conjoined with a desire for future pleasure or an aversion to future pain. When it is conjoined with a preference for an immediate lesser good—and, perhaps, the strength of mind needed to overcome violent fears of missing out on the greater pleasure—the belief that an action will yield pleasure won't be followed by the action's performance. For further defence of the traditional reading see Jonathan Harrison, *Hume's Moral Epistemology* (Oxford: Clarendon Press, 1976), 5–6, and Elizabeth S. Radcliffe, "Hume on the Generation of Motives: Why Beliefs Alone Never Motivate," *Hume Studies* 25.1–2 (1999): 101–22.

3 Hume defines the understanding as "the general and more established properties of the imagination" (T 1.4.7.7; SBN 267), and thinks of it as the faculty that generates our core causal beliefs; see too T 1.3.9.19n (SBN 117) and T 2.3.9.10 (SBN 440). The arguments at T 2.3.3 make it clear that "reason" is being used to denote the source of both our demonstrative knowledge and causal beliefs.

4 On the inference from behavior to motive see, e.g., T 3.2.1.1 (quoted in the text below), 2.3.2.6, 3.1.2.3–4, 3.2.1.2–7, and 3.3.1.4 (SBN 477, 410, 471, 477–9, 575). Though some traits such as wit and social ease are classified as virtues because they are "immediately agreeable" (T 3.3.1.28; SBN 590), others are known to be such because they are useful to their possessor, or useful to others; see T 3.3.1.30 (SBN 590) and EPM 5–8 (SBN 212–67). Though the understanding plays a minimal role (if any) in determining the virtue of characteristics immediately agreeable to their possessors (T 3.3.1.27; SBN 589), it plays a substantial role in classifying as virtues traits that are agreeable to others (T 3.3.1.7; SBN 575), and it is clearly needed to determine what is useful. On the inference necessary to connect motives to utility see EPM App.1 (quoted in the text below), T 3.3.1.9, 3.3.1.19–30, and 3.3.3.9 (SBN 285–94, 577, 584, 606). In the *Treatise*, Hume's explanatory ambitions lead him to emphasize *sympathy*: a mechanism that

explains how we come to hate what we know to be harmful, useless, and ugly, and love what is helpful, useful, and beautiful. In EPM, Hume has either abandoned his aim of offering a reductive account of the phenomenon, or he has decided that his hypotheses on this matter are not suitable for a work intended for a general audience. In the latter work Hume therefore uses "humanity" more than "sympathy" and neither is given a mechanistic explication. L. A. Selby-Bigge effectively argues for this change in emphasis in his introduction to the *Enquiries* (Oxford: Clarendon Press, 1893/1975), xxvi. See, in particular, the explanatory humility Hume expresses at EPM 5.1.2n1 (SBN 213).

5 Christine Korsgaard, "Skepticism about Practical Reason," *Journal of Philosophy* 83.1 (1986): 5–25, and Elijah Millgram, "Was Hume a Humean?" *Hume Studies* 21.1 (1995): 75–93. See, too, Annette C. Baier, *A Progress of Sentiments* (Cambridge, MA: Harvard University Press, 1991); Jean Hampton "Does Hume Have an Instrumental Conception of Practical Reason?" *Hume Studies* 21.1 (1995): 57–74; and Derek Parfit, "Reasons and Motivation," *Proceedings of the Aristotelian Society, Supplementary Volume* 71 (1997): 99–130. Korsgaard says that Hume "presupposes" an answer to questions about "what reason is," and she equates this with his presupposing an answer to questions about what sorts of operation and judgment are "rational." Millgram claims that both of Hume's arguments are "certainly question-begging" (80).

6 In fact Korsgaard doesn't just assign Hume's motivational skepticism to content skepticism; she claims that *all* motivational skepticism must be grounded in content skepticism. ("Skepticism," 6)

7 On the reading I will defend, the claim that passions are necessary for moral and prudential belief is a premise in the first, but not the second of Hume's two arguments; and though the claim that reason can be disassociated from empathy and sympathy is a premise of Hume's second argument, it is not wholly unmotivated, but is instead supported by considerations already advanced in the *Treatise*.

8 See T 2.3.3.5, T 3.1.1.9, and App.4 (SBN 415, 458, 625).

9 These characterizations line up with Millgram's rather than Korsgaard's usage, as Korsgaard treats instrumentalism as a form of skepticism ("Skepticism," 6), and Millgram is concerned to argue that Hume is what we are calling a "full skeptic" rather than an instrumentalist (78–9, and 90n11). This choice in terminology won't have a substantive impact on the discussion to follow.

10 Though Hume felt comfortable using "the moral sense" to denote the set of faculties deployed in central cases of moral judgment (i.e., empathetic imagination integrated with sympathy or humanity) there are significant differences between his account and those of the other sentimentalists. These differences are drastic enough to lead some commentators to deny that Hume's really is a moral sense theory. See, e.g., Jacqueline Taylor, "Justice and the Foundations of Social Morality in Hume's *Treatise*," *Hume Studies* 24.1 (1998): 5–30. In point of fact, Shaftesbury's chief motivation for calling the moral faculty a sense is the *automaticity* with which it is engaged. "Now as in the sensible kind of objects, the species or images of bodies, colours, and sounds, are perpetually moving before our eyes and acting on our senses, even when we sleep; so in the moral and intellectual kind, the forms and images of things are no less active and incumbent on the mind, at all seasons, and even when the real objects are absent. In these vagrant characters or pictures of manners, which the mind of necessity figures to itself, and

carries still about with it, the heart cannot possibly remain neutral; but constantly takes part one way or other. However false or corrupt it be within itself, it finds the difference, as to beauty or comeliness, between one heart and another, one turn of affection, one behaviour, one sentiment and another; and accordingly, in all disinterested cases, must approve in some measure what is natural and honest, and disapprove what is dishonest and corrupt," *An Enquiry Concerning Virtue or Merit* (1699), 1.2.3; reprinted in his *Characteristics of Men, Manners, Opinions, Times*, ed. Lawrence Klein (Cambridge: Cambridge University Press, 1711/1999), 172. And Hume joins Shaftesbury in thinking that, at least over the long run, a normal person cannot help considering the motives behind actions and "in all disinterested cases" feeling appropriate sentiments for the affected parties. The chief difference between Hume and prior moral sense theorists is that (at least in the *Treatise*) Hume sets out to offer a philosophical explanation of how the moral sense works. ("Were nothing esteem'd virtue but what were beneficial to society, I am perswaded, that *the foregoing explication of the moral sense* ought still be received" (T 3.3.1.25; SBN 588; emphasis added.) Hume thinks that such an explanation is possible because unlike colors, feelings of warmth, and the like, the sentiments generated by the moral sense are *secondary* impressions so that ideas to which we have introspective access are involved in their etiology and operation (T 1.1.2.1; SBN 7). Taylor is right to insist, however, that on Hume's view the understanding plays a role in moral judgment far beyond the one it plays in perception, and our moral judgments are consequently more affected by learning and socialization than are our perceptual judgments. See, for example, EPM 9.1.8n1 (SBN 274), quoted in note 46 below.

11 See Clarke's *Discourse Concerning the Unchangeable Obligations of Natural Religion* (1706); reprinted, in part, in *The British Moralists: 1650–1800*, ed. D. D. Raphael, 2 vols. (Indianapolis: Hackett Press, 1991), 1:§224–61.

12 *Groundwork for the Metaphysics of Morals*, ed. and trans. Allen Wood (New Haven: Yale University Press, 1785/2002), 5.

13 To be fair, Hume makes a limited case for (ii) when discussing hedonistic calculation at T 2.3.3.3 (SBN 414). He there provides intuitive support for the claims: (a) that the understanding can generate the judgment that more overall pain than pleasure will result from a given course of action, but (b) that in the absence of a metaphysically distinct aversion to one's greater pain, this understanding will not lead to the action's forbearance. But Hume doesn't there explain whether in judging that some action will bring me more long-term pain than pleasure I *therein* judge that the action is imprudent, nor does he consider the judgment that imprudence is a vice, and the judgment that I oughtn't indulge my vices. Are these judgments ever generated by the understanding alone? Do they ever motivate prudential action? The reader is forced to wait until Book 3 for answers to these questions—answers I will sketch below.

14 Cf. T 3.1.1.9 (SBN 458) where Hume uses the non-representational function of the passions to argue that moral distinctions are not founded on reason and T App.4 (SBN 625) where Hume contrasts our beliefs with our passions by arguing against the view that belief "does not modify the conception but is only annex'd to it, after the same manner that will and desire are annex'd to particular conceptions of good and pleasure."

15 Those theorists of recent vintage who have argued that intentions are a special kind of belief would dispute this reasoning. See, for example, Kieran Setiya, *Reasons*

without Rationalism (Princeton: Princeton University Press, 2007), and David Velleman, *Practical Reflection* (Princeton: Princeton University Press, 1989). But Hume's failure to supply independent arguments against contemporary accounts of this sort cannot be legitimately construed as begging the question.

16 *The Possibility of Altruism* (Princeton: Princeton University Press, 1970), 80n. For a theoretical development of similar ideas see Patricia Greenspan, *Emotions and Reasons: An Enquiry into Emotional Justification* (New York: Routledge Press, 1988).

17 For what I take to be a compelling argument against the truth-directed view of desire (an argument that applies to other kinds of passions as well) see Dennis W. Stampe, "The Authority of Desire," *Philosophical Review* 86.3 (1987): 335–81, 338–9.

18 Compare with Setiya's, "Hume on Practical Reason," *Philosophical Perspectives* 18 (2004): 365–89, 370.

19 *Critique of Pure Reason*, trans. Norman Kemp Smith (New York: St. Martin's Press, 1781/1929), A126, 147.

20 See T 1.3.4.1, 1.3.4.3, 1.3.5.1, 1.3.6.2, 1.3.6.6, 1.3.6.15, 1.3.8.2, 1.3.8.5, 1.3.8.8, 1.3.8.10, 1.3.8.11; EHU 4.1.12 and EHU 4.2.1–2 (SBN 82, 84, 87, 89, 93, 98, 100, 101, 102, 103; SBN 30–1 and SBN 32–3).

21 T 1.3.15 (SBN 173–5). Hume fails to draw the contemporary distinction between rules of inference (such as modus ponens) and belief-forming norms that might be extracted from these rules. Indeed, though some of the rules he supplies are expressed in explicitly normative language stating what we "must" believe in a certain situation, the initial "rules" are phrased as mere generalizations about what must be the case for an object or event to constitute the cause or effect of another.

22 See too Hume's claim at T 1.3.16 (SBN 176) that non-human animals cannot run through arguments but that they reason and infer just as we do when we form our core empirical beliefs.

23 Hume is quite clear that the operations of the passions are just as "rule-conforming" as are the operations of the understanding. So it is clear that an operation's merely conforming to rules—rather than its being influenced by a representation of those rules—does not distinguish reason on either Hume's or Kant's conception of the faculty. See, for instance, Hume's claim, "If snow or flame be presented anew to the senses, the mind is carried by custom to expect heat or cold, and to *believe*, that such a quality does exist, and will discover itself upon a nearer approach. This belief is the necessary result of placing the mind in such circumstances. It is an operation of the soul, when we are so situated, as unavoidable as to feel the passion of love, when we receive benefits; or hatred, when we meet with injuries" (EHU 5.1.8; SBN 46).

24 I have in mind here Hume's distinction between memories and perfect ideas where the latter have "entirely" lost their vivacity (T 1.1.3.1; SBN 8); the division between being angry and merely thinking of anger drawn in terms of the "actuation" of a man (EHU 2.2; SBN 17); Hume's claim at T 2.2.8.4 (SBN 360) that we must use "careful and exact experiments, to separate and distinguish" emotions from ideas because "custom" often makes us "insensible" of the former; his assertion, when accounting for pity and malice at T 2.2.9.2 (SBN 381), "that 'tis not the present sensation alone . . . which de-

termines the character of any passion, but the whole bent or tendency of it," and the frustration Hume expresses with his efforts to provide a phenomenological account of belief at T 1.3.7 (SBN 94) and T App.2–9 (SBN 623–7). Of course, despite his rejection of phenomenology at T 2.3.3.7–8 (SBN 416–7), Hume goes on to employ phenomenologically grounded distinctions later on in the *Treatise* and *Enquiries*, so his dissatisfaction with the methodology was evidently incomplete.

25 See, on this score, Hume's sharp criticism of Wollaston's equation of immoral action with misrepresentation at T 3.1.1.15n (SBN 461).

26 Cf. T 1.4.4.1 (SBN 225) following Hobbes, "Man's *nature* is the sum of his natural faculties and powers. . . . Of the powers of the mind there be two sorts, cognitive, imaginative or conceptive—and motive . . . as appetite is the beginning of animal motions towards something that pleaseth us; so is the attaining thereof the *end* of that motion, which we also call *scope*, and aim, and final cause of the same," *Human Nature: or the Fundamental Elements of Polity*, 2nd ed. (1651) in *The English Works of Thomas Hobbes of Malmsesbury*, ed. John Bohn, vol. 4 (London: Adamant, 1839/2005), 1.4–9; reprinted in Raphael (*British Moralists*, §1). Note that Kant joins Hume in rejecting the idea that we might simply assign reason a conjunctive function, and this motivates Kant to try to demonstrate the single, unified mark of reason in both its theoretical and practical applications. "I require for a critique of a pure practical reason that if it is to be completed, its unity with the speculative in a common principle must at the same time be exhibited, because it can in the end be only one and the same reason that is distinguished in its application" (Ak 4:391; *Critique of Pure Reason*, 7).

27 Bricke, *Mind and Morality: An Examination of Hume's Moral Psychology* (Oxford: Clarendon Press, 1996), 25; G. E. M. Anscombe, *Intention*, 2nd ed. (Cambridge: Harvard University Press, 1957/2000). See too Setiya (*Practical Reason*, 374–5). I think Bricke does Hume a disservice by saddling him with a distinction in direction of fit between desires and beliefs (rather than the faculties that generate them). Because Hume adopts the ontology of a homuncular functionalist he is able to avoid the criticisms faced by those who flesh out the metaphor in crudely behaviorist ways. See, e.g., Smith's behaviorist analysis in *The Moral Problem* (Cambridge: Cambridge University Press, 1994) and Setiya's criticism of Smith in "Explaining Action," *Philosophical Review* 112.3 (2003): 339–93, 364–6.

28 Famously, Hume then goes on to argue that thinking of reason as the fallible cause of our demonstrative knowledge undermines the certainty of our demonstrative beliefs, where a contingent limitation in imagination is the only thing keeping us from wholesale skepticism.

29 *Essays on the Active Powers of Man* (1788), 3.3.1.6–7; reprinted in Raphael (*British Moralists*, vol. 2, §857).

30 Cf. Harrison, *Moral Epistemology*, 7–9 and 20–1.

31 On Hume's claim that we cannot be obligated to act if there is no substantive reason for us to act as we are so obligated see T 3.2.5.4 (SBN 517) quoted in the text below, and Hume's argument that supposing "that the mere regard to the virtue of the action, may be the first motive, which produc'd the action, and render'd it virtuous, is to reason in a circle" (T 3.2.1.4; SBN 478).

32 The "Hobbesian" account I have sketched is perhaps not quite Hobbes's own as I know of no place where he equates the finality of an appetite with its felt strength. Moreover, it is possible that Hobbes assigns a role to what he calls "counsel" beyond the one I have described in the text. See Hobbes's comparison of the imagined consequences of a prospective action to arguments that *spur* deliberation. (*The English Works*, 4.13.5; Raphael, *British Moralists*, vol. 1, §20)

33 See T 1.3.10.3 (SBN 119) and T 2.3.4.1 (SBN 418).

34 The point is reiterated at T 2.3.8.13 (SBN 437).

35 "As repeated custom and its own force have made everything yield to [calm passion], it directs the actions and conduct without that opposition and emotion, which so naturally attend every momentary gust of passion" (T 2.3.4.1; SBN 418).

36 See, e.g., T 3.2.1.8 (SBN 479)—quoted in the text below—and EPM App.4.3–4 (SBN 314–5).

37 For interpretations of Hume on critical evaluation that are similar to the one I will be defending in one respect or another see Don Garrett, *Cognition and Commitment in Hume's Philosophy* (Oxford: Oxford University Press, 1997); David Fate Norton, *David Hume: Common-Sense Moralist, Sceptical Metaphysician* (Princeton: Princeton University Press, 1982); and Corliss Swain, "Passionate Objectivity," *Nous* 26.4 (1992): 465–90.

38 He somewhat less plausibly maintains that if we are convinced that the action does not reflect the agent's character we won't blame him at all (T 2.3.2.6; SBN 410). While Hume might try to defend this as a prescriptive claim on the basis of his views regarding personal identity, it surely does not describe common practice. Still, for arguments against traits of character that take Hume quite seriously here see John Doris, *Lack of Character: Personality and Moral Behavior* (Cambridge: Cambridge University Press, 2002).

39 See too T 2.3.2.6, 3.1.2.3–4, 3.2.1.2–7, and 3.3.1.4 (SBN 410, 471, 477–9, and 575).

40 See too T 3.3.1.9 (SBN 577), T 3.3.1.19–29 (SBN 584–90), T 3.3.3.9 (SBN 606), EPM.5–8 (SBN 212–67) and, "And 'tho such interests and pleasures touch us more faintly than our own, yet being more constant and universal, they counter-balance the latter even in practice; and are alone admitted in speculation as the standard of virtue and morality. They alone produce that particular feeling or sentiment, on which moral distinctions depend" (T 3.3.1.30; SBN 590).

41 Hume emphasizes the need to use reason in the first sort case at EPM App.1.2 (SBN 285), "One principal foundation of moral praise being supposed to lie in the usefulness of any quality or action, it is evident that reason must enter for a considerable share in all decisions of this kind; since nothing but that faculty can instruct us in the tendency of qualities and actions, and point out their beneficial consequences to society and to their possessor."

42 This view of the relation between sympathy and the merit of justice is clearest in EPM, but it is expressed in the *Treatise* too. See, e.g., "Sympathy is a very powerful principle in human nature. . . . We find, that it has force sufficient to give us the strongest sentiments of approbation, when it operates alone, without the concurrence of any other principles; as in the cases of justice, allegiance, chastity and good manners. . . . Justice is certainly approv'd of for no other reason, than because it has a tendency to the

public good: And the public good is indifferent to us, except so far as sympathy interests us in it. We may presume the like with regard to all the other virtues, which have a like tendency to the public good. They must derive their merit from our sympathy with those who reap any advantage from them: As the virtues, which have a tendency to the good of the person possess'd of them, derive their merit from our sympathy with him" (T 3.3.6.1; SBN 618). Hume's account of the artificial virtues is exceedingly complex and (at least in the *Treatise*) quite clearly involves an element of skepticism. For discussions with which I largely agree see David Gauthier, "David Hume, Contractarian," *Philosophical Review* 88.1 (1979): 3–38, and "Artificial Virtues and the Sensible Knave," *Hume Studies* 18.2 (1992): 231–53.

43 For Hume's complicated account of how sympathy operates to convert the imagined suffering of others into moral disapprobation or disapproval see T 2.1.11, 2.2.5, 2.3.6, 2.3.9, and 3.3.2 (SBN 316–24, 357–65, 424–7, 438–48, and 592–601).

44 See T 3.3.6.3 (SBN 619). Hume's dismissive attitude toward moral skepticism is best expressed at the outset of EPM.

45 See too, "It seldom happens that we do not think an enemy vicious, and can distinguish betwixt his opposition to our interest and real villainy or baseness. But this hinders not, but that the sentiments are, in themselves, distinct; and a man of temper and judgment may preserve himself from these illusions" (T 3.1.2.4; SBN 471).

46 Hume emphasizes that correction of the sentiments requires reason when he compares an "untaught savage" who "regulates chiefly his love and hatred by the ideas of private utility and injury" to those of us "accustomed to society" and to "more enlarged reflections" who can by certain "suppositions and views . . . correct, in some measure, our ruder and narrower passions" (EPM 9.1.8n1; SBN 274–5). See too EPM 5.2.22–7 (SBN 225–6).

47 For an application of recent research on acquired psychopathy to the philosophical debate between "internalist" and "externalist" conceptions of the connection between moral judgment and moral motivation see Adina Roskies, "Are Ethical Judgments Intrinsically Motivational? Lessons from 'Acquired Sociopathy,'" *Philosophical Psychology* 16.1 (2003): 51–66. Mightn't Hume think that the psychopath's memory of his moral sentiments must retain some trace of the motivational force possessed by the impression it represents? Not if that memory can degenerate into what Hume calls a "perfect idea" which has "entirely" lost its vivacity (T 1.1.3.1; SBN 8). Though the evidence is hardly dispositive, I think Garrett (187–204) argues in a fairly persuasive way that Hume thinks moral judgments can be made in the absence of their corresponding impressions. At any rate, as I insist in the text, Hume's positive theory of moral judgment should have led him to admit this possibility.

48 For an interesting discussion of the case see Antonio Damasio, *Descartes' Error* (New York: Avon Press, 1994). In this work and others, Damasio argues that reason and the emotions are realized in overlapping anatomical structures. His primary evidence for this is that emotional deficits undercut the operation of faculties needed for prudence. As should be clear from the text, Hume has an explanation of this observation that is compatible with a sharp distinction between reason and passion; it is the kind of explanation given by Patricia Greenspan in her "Responsible Psychopaths," *Philosophical Psychology* 16.3 (2003): 417–29.

49 See R. M. Hare, *The Language of Morals* (Oxford: Oxford University Press, 1952/1991). I thank an anonymous referee for *Hume Studies* for pressing me on this issue.

50 Relevant here is Shaun Nichols's fairly recent study showing that ordinary speakers think psychopaths can know that what they are doing is wrong. See Nicholas's "How Psychopaths Threaten Moral Rationalism: Is it Irrational to be Immoral?" *The Monist* 85.2 (2002): 285–304. The intuitive nature of this claim is brought out when we compare moral judgments to judgments of color—a comparison to which Hume was himself prone. Even on Hume's radically empiricist view of the matter, someone who loses color vision late in life does not therein lose his color concepts or ideas of color. And though the subject of such damage might infer that color-sighted people are now mistakenly calling the things around him "red," "blue" and the like, it is much more likely that he will regard himself as being deprived of his previously held ability to directly identify the colors. Similarly, it would be odd to suggest that someone who loses sympathy late in life therein loses his moral concepts or ideas of virtue and vice. And though the late-onset psychopath might infer that the sympathetic people around him are now mistakenly calling various actions "virtuous" and others "vicious," it is much more likely that he will regard himself as having been robbed of the ability to directly identify these qualities. (He may not care that he has lost this ability, but that is another matter.) Of course, if Hume is right, when someone loses all sympathy or humanity she will have lost her capacity to control those passions that incline her to the acts that she (still) knows are vicious. And our judgments of blame and responsibility will have to take this fact into account.

51 Still, the differences between Hume and Kant should not be overestimated. For while Kant thinks a good will meeting this description must be possible, he allows that, for all we know, it has never actually played a role in any actual action (Ak 4: 407; *Groundwork*, 22–3). Kant's discussion of how reason might give rise to the requisite motive—i.e., esteem or respect for the moral law—is notoriously obscure and intriguing; see Ak 4:401 (*Groundwork*, 17); and Ak 5:71–89; ed. and trans. Mary Gregor, *Critique of Practical Reason* (Cambridge: Cambridge University Press, 1788/1997), 62–75.

52 Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980) tries to assimilate the second use of "reason" to the third, and the third to the first; Jonathan Dancy, *Practical Reality* (Oxford: Oxford University Press, 2000) tries to execute a similar strategy in the opposite direction. For the purposes of this discussion we can remain neutral as to whether any such project will prove successful. Again, the issue here is whether Hume denies that there are any practical reasons in the second sense of "reason" given.

53 For the relevant distinction between "internal" and "external" reasons see Bernard Williams, "Internal and External Reasons," in his *Moral Luck* (New York: Cambridge University Press, 1981), 101–13. The response to Williams's arguments that I think is best supported by Hume's account of moral judgment is much like that supplied by John McDowell in "Are Moral Requirements Hypothetical Imperatives?" and "Might There Be External Reasons?" reprinted in his *Mind, Value and Reality* (Cambridge, MA: Harvard University Press, 1998), 77–111. Speaking of a radically immoral person, McDowell says, "The transition to being motivated [morally] is a transition *to* deliberating correctly, not one effected *by* deliberating correctly; effecting the transition may need some non-rational alteration such as conversion" (*External Reasons*, 107). The difference

is that because McDowell joins neo-Aristotelians in assigning the necessary emotional sensitivity to reason, he rejects Hume's division between reason and passion.

54 At least Hume would say as much. Of course, the man could still be appropriately *attributed* thoughts about color (as when we describe him as wondering what red things look like), but Hume's empiricism has him denying that the man could really have an idea of color. I express this in the text as the view that the man would lack an "adequate grasp" of color concepts.

55 For Hume's comparison of impressions of color and the sentiments of approbation and disapprobation he thinks essential to virtue and vice see T 3.1.1.26 (SBN 468). In some work, Korsgaard focuses on the reasons one has for adopting the general point of view rather than the reasons one has to perform actions approved of from that perspective. See Korsgaard, "The General Point of View: Love and Moral Approval in Hume's Ethics," *Hume Studies* 24.1–2 (1999): 5–30. If Hume were a "realist" he could say that unless one considers situations with the requisite neutrality, one won't be able to figure out which motives are virtuous and so will likely fail to do what one ought. But, Korsgaard argues (*General*, 4–5 and 22–3), Hume's anti-realism prevents him from offering these (decidedly moral) reasons for thinking of morally loaded situations with a neutral eye. The argument does not work. As I argue in the text, even if I think something's being colored consists in its giving rise to visual experiences of a certain sort, I can still think that a person should have surgery to repair his color-blindness so that he may experience and directly discriminate these colors. Similarly, Hume might say that one should adopt the general point of view so as to experience the sentiments that allow one to directly discriminate virtue from vice and therein figure out how to act. Nevertheless, as Korsgaard points out, since Hume is not primarily concerned with the proper attitude to moral skeptics, but, instead, with the historical development of our moral concepts and judgments, he does not dwell on these distinctively moral reasons for adopting the general point of view, but instead provides a causal explanation that gives central place to our deep need to communicate with one another and arrive at shared opinions.

56 See T 3.3.1.26 (SBN 589). Thus Hampton is wrong to suggest that Hume will only criticize in the language of virtue (and will not adopt the language of reason) because he does not think that the traits in question are bad from the agent's own point of view. See her *The Authority of Reason* (Cambridge: Cambridge University Press, 1998): 142–51. On the other hand, even if we diverge from Hume's preferred methodology by trying to come up with as coherent an analysis of "reasonable" as can be reconciled with ordinary usage, we won't wind up making first-person acknowledgment of a motivational problem a necessary condition on its constituting genuine unreasonableness. After all, the kind of extreme pathology that marks psychopathy is often sincerely disavowed.

57 EPM 9.2 begins, "Having explained the moral approbation attending merit or virtue, there remains nothing but briefly to consider our interested obligation to it, and inquire whether every man who has any regard to his own happiness, and welfare, will not best find his account in the practice of every moral duty." Furthermore, the beginning of EPM 9.2.9 (SBN 282) explicitly inquires into what reasons (or "pretext") there may be for preferring vice to virtue "with a view to self-interest." The violation of property rights (an artificial vice) is then singled out as the one possible case in which immorality might sometimes better serve self-interest than does virtue.

58 In Hume's view this is precisely our situation with regard to non-human animals—our relations with them are governed by humanity but not justice; see EPM 3.1.18–9 (SBN 190–1).

59 Though in the final two paragraphs of the section Hume suggests that the risks of cheating are rarely worth it, and that the ends acquired through deceit hardly compensate for the peace of mind that must be sacrificed to attain them.

60 For criticism of this idea see Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge, MA: Harvard University Press, 1985), 5–6.

61 As Kant famously states, "Whoever wills the end, also wills (insofar as reason has decisive influence on his actions) the means that are indispensably necessary to it that are in his control. As far as volition is concerned, this proposition is analytic" (Ak 4:417; *Groundwork*, 34).

62 This is essentially the position adopted by James Dreier, "Humean Doubts about the Practical Justification of Morality," in *Ethics and Practical Reason*, ed.Garrett Cullity and Berys Gaut (Oxford: Oxford University Press, 1997), 81–100. (Dreier makes it clear that his use of "Humean" is meant to be neutral as to whether the instrumentalist view was Hume's own.)