



What's Special about Humeanism

Author(s): Donald C. Hubin

Source: *Noûs*, Vol. 33, No. 1 (Mar., 1999), pp. 30-45

Published by: Blackwell Publishing

Stable URL: <http://www.jstor.org/stable/2672014>

Accessed: 15/12/2008 16:34

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



Blackwell Publishing is collaborating with JSTOR to digitize, preserve and extend access to *Noûs*.

<http://www.jstor.org>

*What's Special About Humeanism**

DONALD C. HUBIN
The Ohio State University

The most common way of classifying theories of reasons for acting begins by distinguishing those that are Humean from those that are not. Humean theories of practical rationality are not *Hume's* theory—he seems not to have anything that could be called a theory of *practical* rationality.¹ But they are descendants of his theory and they carry on the spirit of the Humean critique of those who treat the faculty of reason as a source of reasons for acting. The hallmark of what are now called Humean theories of practical rationality is the assertion that “reason is... the slave of the passions”²—that reasons are based, ultimately, on subjective, contingent, conative states of the agent. Anti-Humean theories deny this, and, apart from this negative thesis and its implications, there is little else that they have in common.

Humeanism, it is fair to say, is the theory to beat; perhaps it is even accurate to think of it as the default position (Nozick, 1993, p. 133). Some form of Humeanism *seems* to underlie the theories of rationality that predominate in the social and behavioral sciences as well as those employed by decision- and game-theorists.³ And, as Robert Nozick points out, most anti-Humean theories incorporate Humean reasons as a *part* of the story about practical rationality. Usually, anti-Humeans admit that Humean considerations are *typically* reasons, even if there are held to be extra-Humean sources of reasons or constraints on Humean reasons. Even those critics who reject Humean reasons altogether must offer a theory with significant *extensional* overlap with the Humean theory—so plausible are many of the implications of that theory.⁴ It is probably safe, then, to say that the burden of proof is on the anti-Humean to move us away from the Humean position. At least, it seems that most anti-Humeans have argued as if this were true—accepting the burden as appropriate.

Is this situation warranted? Apart from its popularity in the social and behavioral sciences—which, after all, we might suspect is not based on a serious appraisal of its philosophical adequacy—is there anything about the Humean position

to justify this favored status? I think there is, though it hasn't been formulated in a way that makes it as compelling as one would like. In fact, I think that what is special about the Humean position on reasons for acting is approximately what most defenders and detractors alike are prone to point to as its attraction. What I intend to do here is argue that, contrary to what the detractors might urge, what appears to be attractive about Humeanism really is a virtue of the theory. I believe that it is a virtue shared by none of its competitors, but I cannot argue that here. Because of this, and because I do not defend Humean theories against other criticisms that have been raised against them, what I say amounts to far less than a wholesale defense of the Humean position. It is, rather, a defense against a particular attack made on it—an attack that has sought to show that Humeanism did not even deliver its advertised goods.

What Appears to Be Special About Humeanism

What attracts many of us, to the different degrees that we are attracted, to Humeanism is, as many have suggested, a motivational argument.⁵ Humean theories appear to connect reasons for acting with motivation in an appealing way. Reasons for action should turn out to be the sort of things that typically *do* motivate people. The appeal of a connection between an agent's reasons for acting and her motivation is shown by the fact that many anti-Humeans aim in their own theories of practical rationality to tie an agent's reasons to her motivation. The way the Humean secures this connection is obvious: reasons are connected to motivation because they are grounded in the agent's desires (values, concerns, *etc.*), and these conative psychological states are, *inter alia*, motivating states.

As Michael Smith points out (1994, pp. 136-7), though, a theory can make the connection between an agent's reasons for acting and her motivation too tight, as well. We want to account for why reasons motivate people without our explanation entailing that they always do so. The obvious problem for anti-Humean theories is making the connection between reasons and motivation; but Humeans must take care not to make the connection so close that they can't allow for divergence between reasons and motivation.

Were we to require that reasons for acting always motivate, we would make the connection between reasons and motivation too tight. It seems apparent that people sometimes lack any motivation to perform actions they have the most reason to perform—even actions they *judge* they have the most reason to perform. The possibility of true, practical irrationality seems undeniable. And if this is so, the Humean cannot claim that those considerations recognized by his theory as being reasons for action will always motivate an agent. It appears that the best the Humean can do to link reasons with motivation is to say that "*insofar as an agent is practically rational, she will be motivated to do what she has reason to do*".

Unfortunately for the Humean, this is hardly a distinguishing virtue of the theory for, as Christine Korsgaard points out (1986, p. 23), every other theory of

practical rationality can say the same.⁶ Consider a theory that holds that a person has a reason to act just in case the action would, from some hypothetical standpoint, be approved of by everyone. There is no guarantee that real people *will* be motivated by such considerations. But, being motivated by the reasons one has to act is part of what it is to be ideally rational. So, we can assert with confidence that *were* one practically rational in the appropriate sense, one would be motivated by these considerations.

It appears that no plausible theory can boast that people will necessarily be motivated by those considerations it recognizes as reasons—an improvident boast anyway in light of the possibility of true practical irrationality. And, every theory can boast that insofar as people are rational (according to that theory's account) they will necessarily be motivated by those things that are deemed reasons for acting by that theory—an empty boast. The special connection to motivation that Humeans tout for their theory of practical reasons seems absent. So, what's special about Humeanism? To answer this question, we will take Humeanism apart.

Decomposing Humeanism

Humeanism can, I think, be profitably decomposed into two component theories. The first I call, following Shelly Kagan (1989), 'pure instrumentalism'. The second I follow Stephen Darwall (1983) in calling, 'the thesis of desire-based-reasons'. But I need to make clear that I use 'desire' very broadly to include any positive conative state that might plausibly be claimed to motivate action. In particular, it covers states that a more sensitive psychology would call 'caring about' or 'valuing', and distinguish from desire in a narrower sense.

The thesis of pure instrumentalism holds that reasons are communicated from ends to means—that he who has a reason for the ends has also a reason for the means. For our purposes, the means ends connection must be interpreted in a broad—though, I don't think, idiosyncratic—sense to include not only causal connections but also criterial connections and part whole (mereological) relations. Thus, we will, somewhat artificially, take climbing Mt. Everest to be a means to climbing the tallest mountain in the world not because there are two events which stand in a causal relation but because the climbing of Mt. Everest satisfies the criteria for climbing the tallest mountain in the world. And, we will take buying lottery numbers based on your children's birth dates to be a means to winning the lottery as a result of buying lottery numbers based on your children's birth dates. Pure instrumentalism asserts only that reasons are transmitted across causal, criterial, and mereological connections; it makes no claim about the source of reasons.⁷

The thesis of desire-based-reasons, on the other hand, holds that the ultimate source of reasons for an agent to act, in the sense relevant to rational advisability and to the rational appraisal of agents, is in the subjective, contingent, conative states of that agent. This thesis does not entail pure instrumentalism any more than it is entailed by it. One could hold to the desire-basedness of reasons without

accepting that reasons are communicated from ends to means. We must remind ourselves, also, that this thesis says nothing about the ultimate grounding of *moral* reasons, *political* reasons, *economic* reasons, *prudential* reasons, *etc.* I take these to be practical reasons in the sense that they are considerations from the specified point of view in favor of action—considerations that tend to determine what is required (or recommended) from that point of view. Here, though, we are concerned with the thesis of desire-based-reasons only with respect to the source of reasons in the sense relevant to *rational* advisability and to the *rational* appraisal of agents and their actions.

By decomposing what has come to be called ‘the Humean theory of practical reasons’ into the theses of pure instrumentalism and desire-based-reasons, we can see what the Humean can say about reasons and motivation that other theories of practical reasons have not been able to parrot. In brief, my argument will be that the thesis of pure instrumentalism should be construed in such a way as to be noncontroversial. But, if pure instrumentalism is noncontroversial, then the Humean is in a position to assert a unique connection between those considerations she acknowledges to be reasons for acting and the agent’s motivation. This is the key to what is special about Humeanism.

Pure Instrumentalism: When properly understood, the thesis of pure instrumentalism is, if not indisputable, at least extremely innocuous. To try to show this, I will begin by imagining a critic attacking pure instrumentalism with apparent counterexamples. And, with respect to this tenet of the Humean theory, I intend to take the dogmatic line of asserting that, properly clarified, there *is* a conceptual confusion in conceding that a person has a reason for an end but no reason whatsoever to produce means to that end. Apparent counterexamples are *apparent only*.

To challenge this view, the critic must assert that a person’s having a reason to bring about an end does not entail that he has any reason to engage in the means required to do so. There are several types of cases that might seem to make this claim plausible. Nevertheless, the Humean denies it. Adjudicating this dispute is a bit trickier than it might seem. We must be careful not to commit either of two errors: we must not overreach our intuitions; and we must not build too much into the thesis of pure instrumentalism.

Suppose that you are on a long hike in the desert and have run out of water. Your thoughts turn to a cold draft of your favorite ale and suddenly you see it sitting before your eyes. Certain things are true of you at this point: you come to desire to drink this pint of ale; you would get great pleasure from drinking it; you would avoid the unpleasant experience of a parched throat by drinking it. For one of these reasons, or for some other, let us suppose, it seems plausible to say that you have a reason to drink the ale. At this point, the Devil appears to you and says the ale is yours if you will just give him the soul of your beloved child. Now, giving up the soul of your child is a means of attaining the end that, it seems plausible to say, you have reason to bring about. If you do have a reason to bring about this end, and if reasons are communicated from ends to means, then you

have a reason to give the soul of your child to the Devil. But isn't it manifestly obvious that an agent could have no reason whatsoever to do *this* regardless of what reason the agent has to get the ale? Some think so.⁸

The case is, I believe, interestingly complex. But one feature, especially, makes it doubtful that we should draw any theory-challenging inferences from our intuitive response to it. This case appears to be one in which any reason you have based on the desire for the ale is "swamped" by other reasons you have. So, the defender of pure instrumentalism might say, you do indeed have *a* reason to give up the soul of your child, but you have such weightier reasons for not doing so that it is difficult to see this reason and misleading to talk of it in most contexts. Perhaps it would be inappropriate even to think about having such a reason in this sort of case. It would, I think, be an error to trust our intuitions about the fine points of when it is correct to say that one has a reason. Those intuitions are shaped by pragmatic considerations as well as semantic ones, and these different sorts of considerations are seldom carefully distinguished. And so, while it may be perfectly *true* to say that you have a reason to turn over your child's soul, it may be quite *improper* to say it, or even to think it. The appeal to what we might call 'the "swamped reasons" defense' is a familiar one, of course, but its appropriateness is sometimes forgotten, I believe.

Avoiding the second error requires that we remember that the thesis of pure instrumentalism does not commit one to any theory about the source of reasons or the proper descriptions of "ends". Theories of practical reasons may subscribe to pure instrumentalism and retain complete freedom about how to differentiate between states of affairs for the purposes of determining which states of affairs we have reason to promote and which we do not.

Recognizing this provides for a different reply to the above example—a reply that some might find more compelling than the "swamped reason" defense. And, even if we find, as I do, that the "swamped reason" defense is adequate for the above case, this new reply will be more plausible in other cases.

A theory of reasons for acting is free to subscribe to the thesis of pure instrumentalism and deny that you have *any* reason, even a swamped one, to turn over the soul of your child. This is because the theory is free to deny that you have any reason to bring about the concrete end that would be brought about in this situation. A pure instrumentalist theory of reasons for acting can treat the fact that you would get a drink of your favorite cold ale when your throat is parched as a feature of the end in question that functions in its favor without conceding that you have a reason to bring about every state of affairs that includes it. For the state of affairs that would be brought about is that of your receiving a drink of ale *by* giving up the soul of your child. And the theory can say that no matter how good the drink of ale might be, you have *no* reason to bring about *this* state of affairs.

It is natural to think of each good feature of the outcome as providing a reason for bringing about the outcome. So that, if there is any good feature of the outcome, then there is *some* reason to bring about that outcome. But this is not part of the thesis of pure instrumentalism, which concerns only the transmission of reasons from ends to means. We might call this other thesis *compositionality*. It

holds that reasons for outcomes are composed of reasons for their parts. A pure instrumentalist is free, though, to reject compositionality. She may, if she wishes, take reasons to be, strictly speaking, holistic—applying only to entire world descriptions.⁹ But it is important to see that there is nothing in the thesis of pure instrumentalism that forces one to determine ends this finely. The most plausible development of the Humean position will, I think, take the level of specificity of the ends to be determined by the intrinsic valuations of the agent. For many of us, this may entail the denial of compositionality. For none of us will it entail that ends must be entire world descriptions.

Rather than illustrate this point further with the above example, let me offer two better suited to the point. The first I will introduce with the assumption that it is the desires (values, concerns) of the agent that ultimately ground reasons to act. But this is for purposes of exposition only. Before leaving the case, I will generalize my comments. The second, designed to illustrate a different phenomenon, will be free of this assumption from the outset.

Suppose that I desperately desire the respect of my estranged brother. I have, then (we are supposing), a reason to bring about the state of affairs in which my brother respects me. A generous (and *very* skilled) psychosurgeon, noting my distress, offers at no charge to wrestle my brother to the ground, administer anesthetic, and do whatever surgery is necessary on my brother's brain to make it the case that he respects me. All I have to do is "give him the go". I might deny that I have any reason (even a swamped one) to accept the offer. And if this is true, then it seems the reason for bringing it about that my brother respects me is not a reason for taking the psychosurgeon up on his offer even though that is a means of securing my brother's respect. Isn't this a counterexample to pure instrumentalism?

I think the pure instrumentalist should hold that if we believe that we have no reason to accept the psychosurgeon's offer, it is because, contrary to appearances, we have no reason to bring about the end this would produce. When I say I desire my brother's respect, I don't mean that I desire every state of affairs in which it is true that my brother respects me. Rather I mean to point to a set of outcomes in which my brother respects me that is typical (or otherwise salient), and say of that set that I have a reason to bring about one of those outcomes. When we agreed that there was a reason to bring it about that my brother respected me, we didn't mean that I had a reason cause his respect no matter what else was true in that situation. Reflection on the psychosurgeon's offer might make me realize what I actually desire, and it is completely natural for me to reply, "I don't want my brother's respect *that way!*"¹⁰ My desire might not simply be for my brother's respect, but for any one of a set of typical cases in which I have that respect. These may all have in common the fact that I have his respect as a result of having earned it.¹¹ Thus, the Humean might say, we have no reason to accept the offer precisely because we have no reason to bring about the result that could be achieved by accepting it. *That* result is not desired.

Now, this reply depends in no way on the fact that, in addition to pure instrumentalism, the Humean accepts the thesis that reasons are based on desires. The pure instrumentalist, in general, is free to accept a grounding for reasons that is

sensitive to any feature of the state of affairs, including such things as: how the events were brought about, when events took place, who caused the events, why they were caused, what alternatives were available, what intentions various agents had, *etc.*. To ensure that we have a counterexample to the pure instrumentalist element of a theory of reasons for acting, the outcomes must be described so as to include every feature that the theory of reasons in question takes to be relevant to the desirability of the outcome. Let us call this “the inclusive description” of outcomes, but understand that it is nothing more than a fuller, more accurate description of the end. For the Humean, the descriptions must include everything that affects the agent’s intrinsic evaluation of a state of affairs.¹² If the agent cares about whether what we would typically speak of as the outcome of an action is produced in one way or another, then, for purposes of applying the Humean theory of reasons for acting, these are two distinct outcomes.¹³

Consider a second example—one drawn from the philosophy of punishment. Standard utilitarian theories of punishment hold that “all punishment is mischief; all punishment in itself is evil.... if it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil” (Bentham, 1970, p. 170). The idea is that there is always a reason not to punish, but that reason can be overbalanced by reasons in favor of punishment. This view does not follow from Bentham’s moral instrumentalism alone; rather it is the conjunction of that moral instrumentalism with his (distinctly implausible) hedonistic theory of value that gives Bentham this result. A pure instrumentalist is free to argue that while the sort of pain involved in punishment is *typically* an evil, it is, nevertheless, “an end in itself that the guilty should suffer pain” (Ewing, 1929, p. 13) and that “[t]he state of affairs where the wrongdoer suffers punishment is morally better than one where he does not” (Rawls, 1955, p. 5). A less retributive, but no less retrospective pure instrumentalist, could hold that the suffering of the guilty is, if not a positive good, at least not an evil or not as great an evil as the suffering of the innocent.¹⁴

Someone holding the sort of view just sketched would take the fact that an act caused suffering to be a *prima facie* consideration against the act only in the sense that, knowing nothing else about the situation, it is reasonable to suppose that there is a consideration against the act. A defender of this view would not, as Bentham does, take the suffering caused by an action to be a *pro tanto* consideration against the act.¹⁵ One who holds *this* view is not committed to saying that there must always be a reason not to cause suffering—a reason that must be overcome by other considerations. There may be *no* reason not to cause the suffering of the guilty, or even a positive reason to cause such suffering solely in virtue of their guilt. Thus, a satisfactory justification of the pain caused to the guilty in the process of punishment need not involve overcoming a moral consideration against causing pain. The full description of the outcome in question must include something like ‘causing pain to the guilty’, and the theory we are imagining does not assert a reason for avoiding such pain even if it recognizes, typically, a reason not to cause pain.¹⁶

Using the inclusive description of outcomes, I maintain that there are no plausible counterexamples to the thesis of pure instrumentalism. Reconsider the psycho-surgeon example. For this to provide a counterexample, it would have to be the case that I have a reason to bring it about *by agreeing to let the psycho-surgeon operate on him* that my brother respects me but I don't have a reason to agree to let the psychosurgeon operate on him. This does not sound possible to me. I wouldn't understand what a person was trying to say if he told me that he had a reason to bring-it-about-*by-doing-y-that-x*, but no reason to do *y*. The concept of pure instrumentalism is so central to our conception of practical rationality that, once we avoid trading on insufficient descriptions of outcomes, we could make little sense of a conception of practical rationality that did not include it.

It is worth noting how broad the appeal of pure instrumentalism is. The theories of value that make states of affairs reason-makers need not be agent-neutral. Certainly, to illustrate with the Humean account, an agent's preferences need not be agent-neutral. I may prefer the state of affairs in which an innocent person is killed by someone else to that in which she is killed by *me*. Similarly, we could construct an agent-relative theory for evaluating states of affairs that is independent of the agent's conative states. Such a theory would establish different reason-making states of affairs, and hence, perhaps, conflicting reasons, for different agents. Thus, it might tell *me* that there is a reason for me to make it the case that you, rather than I, kill an innocent person and tell *you* that there is a reason for you to make it the case that I, rather than you, do so. Such is the nature of agent-relative evaluation. The theory of value could also be time-relative. It might say that the situation in which you kill an innocent person today is worse than that in which you kill an innocent person tomorrow. In addition, a theory of value could be sensitive to the intentions of agents, so that two states of affairs would have to be counted as relevantly different no matter what their other similarities if one was brought about intentionally and the other not.

The breadth of the appeal of pure instrumentalism is further demonstrated by the following fact. Kant, no sympathizer with the intuitions behind Humean theories of practical rationality, says in the *Groundwork* that “[w]ho wills the end, wills (so far as reason has decisive influence on his actions) also the means which are indispensably necessary and are in his power” (1956, p. 84-5). While this is not equivalent to the thesis of pure instrumentalism—it is a weaker claim, at least—I think the most plausible interpretation of the Kantian dictum sees it as grounded in a commitment to the thesis of pure instrumentalism. On this view, his argument with the theories that have been developed into Humean theories of practical rationality concerns not pure instrumentalism but the thesis of desire-based reasons.

If I am right that there is nothing about the nature of pure instrumentalism that prevents its being coupled with agent-relative and time-relative theories of value, then the theory has broad appeal indeed. The argument offered here requires only that the thesis of pure instrumentalism is true. I have tried to provide reasons for thinking this is so—largely by way of indicating what it does *not* entail. I, in fact

think that it is a conceptual truth. It is difficult to imagine a plausible theory of reasons for acting that will not incorporate pure instrumentalism; it is difficult to imagine what a person who understood the meaning of ‘reason’ could be wondering about if she were wondering whether pure instrumentalism is true. And I have tried to give some reason for thinking that this is so. However, for my present purposes, it will suffice to establish that pure instrumentalism is (uncontroversially) true.¹⁷

What *Is* Special About Humeanism

Suppose that pure instrumentalism is, as I have suggested, a noncontroversial element of any plausible theory of practical rationality. How does this help us say what is special about Humeanism? I believe that it allows us to salvage something from the argument from motivation despite the fact that it is manifestly psychologically possible for a person to fail to be motivated to produce means necessary to ends that he desires to bring about. If so, then there might be a plausible path from pure instrumentalism to Humeanism—a path closed to other theories.

I will first put the point crudely. According to the Humean, it is the agent’s valuations that generate (through the vehicle of pure instrumentalism) reasons for acting. Let us say that these valuations define the agent’s evaluative point of view. Not all evaluative points of view are defined in this agent-dependent way. We can, and do, evaluate states of affairs by the amount of happiness they contain, the degree to which they satisfy the desires of all agents, the degree to which people get their due, *etc.*. We could construct an instrumentalist theory of reasons for acting that employs any of these evaluative points of view. And on any of these theories (even Humeanism), an agent is logically free to remain unmoved to perform actions that promote the end that is claimed to generate reasons. But it is only on the Humean theory that the agent is not also logically free to remain unmoved by the evaluative point of view in question; for this point of view is defined by what moves the agent. If, then, we assume that pure instrumentalism is true, we have something unique about the way in which the Humean connects reasons for acting with the agent’s motivational structure. There is a special kind of practical defect in those unmoved by Humean reasons. They run afoul of pure instrumentalism; whereas those unmoved by other sorts of reasons may simply not be moved by those features that ground the evaluation of the states of affairs in question.

But this characterization of what is special about the way Humeanism connects reasons and motivation is too crude.¹⁸ At least this is so if it is possible, as it seems to be, for a person to remain unmoved by states of affairs that she intrinsically values (desires, cares about, *etc.*). I want to do full justice to human psychological quirks. I think doing so requires admission that we can value a state of affairs and remain unmoved to bring it about. Thus, if the Humean uses as the basis for reasons the agent’s intrinsic values, it seems, once again, that even *he* can’t guarantee the connection to actual motivation. Aren’t we back where we

began, with every theory able to say that, insofar as one is rational, according to that theory, one is motivated by those considerations that theory deems to be reasons for acting?

There are these differences. First, while I think an agent may remain unmoved by her own values, there is, nonetheless, a conceptual connection between an agent's values and her motivation. Intrinsic valuing is a mental state that is, I think, best understood functionally, in terms of its causal connections to typical stimulus situations, behavioral outputs and other mental states. While we might be convinced by other evidence that an agent values a state of affairs she is not motivated to bring about, we would certainly count the lack of motivation to produce the state of affairs as direct evidence against this thesis. We might say that to intrinsically value a state of affairs is to be in a state that *typically* motivates action to produce the state of affairs.¹⁹ And this connection is not external to the concept of intrinsic valuation. It is not that intrinsic valuing typically motivates in the sense that dropping a glass typically breaks it. It is part of the very concept of intrinsic valuation that it is the sort of thing that typically motivates action.²⁰

The situation in which an agent is unmoved by what she values (desires, cares about, *etc.*) is an anomalous case—not just in the sense that the person in question is psychologically unusual. Those who are completely unmoved by considerations of sympathy are also psychologically unusual, but their motivational set is not anomalous in the way that we are considering. The situation of being unmoved by one's values presents a conceptual oddity, not merely a psychological or evolutionary deviation.

Second, and more importantly, those who are irrational in the Humean sense cannot “escape” rational criticism in the same way that, it appears, those judged irrational by some other theory can. They may be unmoved to perform actions that are means to their ends and even, in the anomalous case, unmoved by the ends themselves. But the charge of Humean irrationality is “unshuggable” in a way that other conceptions of rationality are not. For the agent who is irrational in the Humean sense cannot avoid the motivational *force* of our judgment about him by pleading that he does not *care* about the ends in question. The motivational force is there, even if the motivation is anomalously absent. It is the nature of desires, values, concerns themselves that they typically produce motivation in the agent who has them. When we base reasons on something other than the subjective, contingent, conative states of the agent, it is difficult to see how a parallel claim can be sustained.

If this is correct, then Humeanism connects reasons for acting with psychological states of the agent that are typically, and conceptually, connected to the agent's motivational structure. When such reasons are presented to the agent, he cannot respond that he doesn't *care* about the evaluative perspective that grounds these reasons. If, caring about the ends, he remains unmoved by them, he presents an anomalous case. And, if, moved by the ends, he is nevertheless not moved to perform the actions he recognizes as means to these ends, he is means ends

irrational—a noncontroversial form of irrationality given the truth of pure instrumentalism.

Is this what is special about Humeanism? I suspect that it is; I suspect that competing theories cannot make a similar connection between reasons for action and an agent's motivational structure. Conjectures aside, I do not believe that any other theory *has* established such a connection. In the absence of an argument to establish such a connection in a competing theory, I think we can at least conclude that what is special about Humeanism is that it *has* established such a connection between reasons and motivation; defenders of other theories have yet to show an analogous connection within their theories.

That Humeanism has this feature is not dispositive, of course. Other attacks can be, and have been, launched against the theory. Some have argued that the Humean conception of practical rationality is not properly normative, or that it has counterintuitive consequences. While I do not share these reservations about the theory, nothing I have said here responds to these worries. My project has been confined to responding to those who have suggested that Humean theories couldn't even offer a special connection between an agent's reasons for acting (in the relevant sense) and the agent's motivational structure.

Why It Matters

There comes a point in all disputes of this sort at which one wants to ask what difference it makes. What hinges on how we decide the issue between the Humean and his critics? This is a question the Humean should welcome for reasons I will try to explain. As I see it, one of the most important differences it makes is tied crucially to the force we take reasons, especially moral reasons, to have. Those who reject Humeanism are frequently motivated by the attempt to show that moral reasons are binding on everyone despite the fact that acting on moral reasons is not always instrumental to the satisfaction of one's desires. If so, then the moral reasons must be reasons, binding on all agents, independently of their connection with the agent's conative, and perhaps motivational, structure.

Little is usually said about the concept of bindingness. It is clear what it means for a requirement to be *morally* binding; it means that the requirement is violated on pain of immorality. But, obviously, *moral* bindingness isn't what is meant in this context. Of course moral reasons are *morally* binding; Humeanism presents no challenge to this truism. Those who see Humeanism as a threat to the bindingness of morality mean something else by 'bindingness'; they mean, I think, '*rationally* binding'. To be blunt about it, some critics of Humeanism are looking for a club with which to beat those who ignore moral reasons. They have already beaten them with the club of immorality—to no effect. They want a bigger club; they want the club of irrationality.

But what, exactly, is it about the charge of irrationality that makes it carry more weight (in a certain context) than that of immorality? It is certainly not that being irrational is *morally* worse than being immoral. In the first place, I'm not

sure what sense can be made of this. If irrationality is morally bad, then it just *is* immoral. Secondly, there are many instances of irrationality that are of no moral consequence at all, and sometimes it would even be morally better for certain people to behave irrationally in the pursuit of their ends. Finally, there would be no point in making the charge of irrationality to a person who was insensitive to moral reasons if irrationality turned out to be only a moral failing.

It will do no better to suggest that the charge of irrationality carries more weight than the charge of immorality because it is worse from the evaluative point of view of the *critic*. The fact that the behavior in question doesn't comport with the reasons generated from the critic's evaluative point of view does not necessarily have any special force for the agent criticized.²¹

The charge of irrationality wields the special force it does, I think, because it is made from the agent's evaluative point of view. It is precisely the fact that in making rational evaluations we adopt the agent's normative standpoint that makes the charge of irrationality have special force for the agent. If we are correct and he understands our rational criticism, he is generally motivated to act on our recommendations. We show him the error of his strategies for getting what he ultimately wants. If our analysis is sound and he remains unmoved, he shows a special kind of practical defect—different in kind from that shown by a person who fails to act on reasons because he doesn't care about the evaluative point of view from which they are generated.

I harbor the suspicion that many anti-Humeans want to keep all the special force of Humean rational criticism, but not leave the content of rational recommendations to the contingent conative states of the agents. I do not believe this can be done. If rationality is something that requires one to count the desires and interests of others equally with one's own, for example, or to pursue some other end(s) that the agent may not value, then rationality itself is not binding in the appropriate sense—the immoralist is as free to say, "Who cares about rationality?" as she is to say, "Who cares about morality?" And what have we gained?

When we ask about whether moral reasons are *reasons* for everyone, though, we sometimes have concerns other than bludgeoning immoralists into acting morally. Perhaps we are simply trying to understand or diagnose the immoralist. The Humean takes irrationality to be a special kind of failing—not a failure to adopt a particular evaluative perspective, but a failure to be motivated by the evaluative perspective one adopts. This means that the Humean analysis gives the diagnosis of irrationality a special place—one that seems plausible but also seems impossible on non-Humean analyses. So part of what makes Humeanism special is that Humeanism makes rationality (and irrationality) special.

Perhaps the deepest concern we have when we ask whether moral reasons are reasons for everyone arises when we are engaged in practical deliberation ourselves and wondering if we have *reason* to do what is morally required. And what does the Humean have to say about this? She *could* say that we are wondering about the effects of acting morally or about what desires we have. Perhaps this is sometimes so. But I suspect that there is often something deeper going on than

simple fact-finding about consequences or desires. I think that we are often engaged not in practical reasoning but in a precondition for practical reasoning. And this precondition is not the discovery of a pre-existing conative structure with which to define our evaluative point of view. We are not trying to *discover* our evaluative point of view; we are *creating* it. Just as we *entertain* a hypothesis to draw out its implications and (sometimes) to see the world as we would if we believed it, we at times “try on” normative stances. If I am right, this can be more than a process of gaining knowledge of our values; it is sometimes an act of self-creation.²²

If this is sometimes what is going on, the Humean would distort the situation if she insisted that this was itself a matter of practical deliberation—of determining means to desired ends. It would also be a distortion for her to treat it as an exercise in theoretical reasoning about what the agent already desires—theoretical reasoning that is antecedent to practical deliberation. But this would be no greater distortion than the non-Humean treating this phenomenon as an exercise in theoretical reasoning about what is objectively good or valuable, or the nature of a rational agent—as if her activity were ultimately one of intellectual discovery.

These questions about what difference it makes whether the Humean is correct or not and about what a theory of practical rationality should do for us are large and difficult ones; they will not be answered easily or decisively. But these are the questions that must be answered, I think. For what is at issue is not, as it sometimes appears, the meaning of a word (like ‘rational’ or ‘reason’); what is at issue is the value of a theory.

Notes

*I’m grateful to Shelly Kagan for awakening in me worries about the instrumental conception of practical rationality and for providing insightful criticisms of my position. Arthur Kuflik provided extremely helpful criticisms when this paper was read at the 24th Annual Hume Conference in 1997. I am also grateful to Robert Audi, Jamie Dreier, David Commiskey and an anonymous referee at this journal for assistance in clarifying my arguments. Additional help was generously offered by participants at the Hume Conference and at the Inter-University Center Conference on Ethics and Political Philosophy in Bled, Slovenia (1994).

¹See Darwall [1983, p. 53], Hubin [1991, p. 26-28], and Milgram [1995]. In my earlier work, I refer to the sort of view discussed here as “Neo-Humeanism”, which is a more accurate, but less familiar, term.

²The full quote is, of course, “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” (Hume, 1968, p. 415).

³Stephen Darwall (1983, pp. 62-77) denies that the Humean theory of rationality (which he refers to as the *Desire-Based-Reasons Thesis*) plays a significant part in modern formal theories of rationality. The meat of these theories, he argues, lies in how they connect choice under certainty with choice under risk and uncertainty. These interesting formal theories retain all of their force quite independently of whether the ranking of ultimate outcomes (the formal preference ranking) is induced by the desires of agents or some other factor. (In the latter case, it might be less misleading to call it a *preferability* ranking.)

⁴See, for example, Jonathan Dancy’s recent *Moral Reasons* (1993).

⁵Darwall (1983) identifies it as one of four roots of the Thesis of Desire-Based-Reasons. Kagan ("On Instrumental Reasons", 1989) singles it out as the primary motivation for a Humean theory. Korsgaard's examination of the relation between content skepticism and motivational skepticism is based on the recognition that the apparent link to motivation offered by Humean theories of practical rationality is one of their primary attractions. Of course, some are moved by other considerations. For example, Joseph Heath (1997, pp. 452ff.) identifies as the Humean's most powerful argument the apparently superior ability of the Humean account of practical reasons to offer "a [teleological] rationalizing explanation of actions" in terms of reasons for acting. (Heath goes on to contend that this argument is committed to an unjustifiable foundationalism and, so, does not, in the end, give support the Humean position.)

⁶Actually, Korsgaard claims that "there is probably no moral theory" that is ruled out by the requirement that what it counts as reasons must motivate an agent insofar as he is rational. However, she means to be talking about theories of practical rationality not theories of morality so I take the reference to moral theory to be a slip. Furthermore, her argument that motivational skepticism must be wholly dependent on content skepticism requires her to argue that *in fact* no theory is excluded by this sort of requirement (not just that it is probable that no theory is so excluded).

⁷It is for this latter reason that it is called 'pure instrumentalism'. It is the purely instrumental portion of the Humean, and (I shall argue) any other plausible, theory of reasons for acting. It would be a serious misreading to understand 'pure instrumentalism' as holding that *all* reasons are instrumental—even worse to understand it as holding that all reasons are instruments to the agent's ends.

⁸For instance, Larry Temkin, from whom I borrow this example, and Jonathan Dancy (1993, p. 181).

⁹See G. E. Moore's discussion of organic unities (1954, pp. 27-33).

¹⁰"Someone says to me: 'Shew the children a game.' I teach them gaming with dice, and the other says, 'I didn't mean that sort of game.' Must the exclusion of the game with dice have come before his mind when he gave me the order?" (Wittgenstein, 1953, p. 33^e). Similarly, were I to say that there is some reason to play a game with the children, I needn't mean that there is some reason, perhaps overridden, to play dice. What I have said is compatible with my thinking that there is *no* reason to play *that* sort of game.

¹¹Arthur Kuflik has pointed out to me that the plausibility of this example may turn on *how* the psychosurgeon will change my brother so that he respects me. If what is necessary is to remove from my brother some neuroses which make him incapable of recognizing my accomplishments—which, if recognized, would produce his respect—I may have no objection to his actions. I am imagining a case in which the respect would be produced more directly and not as a result of having "earned it".

¹²It might be thought that this is a distinctly *un*Humean move; it seems to undermine the distinction between ends and means that is crucial to a Humean theory. I disagree. The world perversely refuses to carve *itself* up into ends and means. If there is to be a division, it is not to be found in ontology but in psychology. And, if there is anything central to what is now called 'the Humean Theory of Practical Rationality' it is that agents may intrinsically desire (value) any coherent state of affairs. *Ends*, then, are what the agent values intrinsically; *means* are states of affairs that stand in the proper causal, criterial or mereological relation to ends.

¹³In general, we can say, there are indefinitely many ways of classifying the set of outcomes of any action. For convenience, we can treat outcomes as indistinguishable if they do not differ in intrinsic value. But, for purposes of evaluating a theory of reasons, we certainly may not treat as indistinguishable two outcomes that the theory holds are of different intrinsic value. Thus, if a theory distinguishes between the state of affairs of my brother respecting me as a result of my good works and his respecting me as a result of forced psychosurgery, we must treat these as different ends. To anticipate the next example, if a theory distinguishes between the suffering of the guilty and that of the innocent, we cannot treat as equivalent two states of affair that involve the same suffering but distribute it differently between the guilty and the innocent.

¹⁴Shelly Kagan offers a very clear discussion of these matters (1998, pp. 54-59). For fuller development, see Vallentyne (1988).

¹⁵See M.B.E. Smith (1973) for a clear distinction between these two views which have often misleadingly been couched indiscriminately in terms of *prima facie* reasons. Kagan (*The Limits of Morality*, 1989, p. 17) introduces the term '*pro tanto*' to distinguish the two views.

¹⁶Alternatively, consider a theory that holds that deterrence of crime is a good thing if, but only if, it is brought about without punishing the innocent. Such a theory would distinguish outcomes not only according to their deterrence rates, but also according to whether these rates are achieved without punishing the innocent. This would have to be included in the very descriptions of the outcomes, for achieving deterrence without punishing any who are innocent is, according to this theory, a morally different outcome than achieving deterrence by punishing some who are innocent.

¹⁷ I am grateful to Shelly Kagan for pointing this out to me.

¹⁸While I believe others may have tried earlier, it was Geoff Sayre-McCord who made me see clearly the inadequacy of the characterization.

¹⁹An attractive alternative conception is suggested by John Stuart Mill. In his notes on his father's *Analysis of the Phenomena of the Human Mind*, the younger Mill says: "I believe the fact to be that Desire is not Expectation, but is more than the idea of the pleasure desired, being, in truth, the initiatory stage of the Will. In what we call Desire there is, I think, always included a positive stimulation to action; either to a definite course of action which would lead to our obtaining the pleasure, or to a general restlessness and vague seeking after it. The stimulation may fall short of actually producing action: even when it prompts to a definite act, it may be repressed by a stronger motive, or by knowledge that the pleasure is not within our present reach, nor can be brought nearer to us by any present action of our own. Still, there is, I think, always, the sense of a tendency to action, in the direction of pursuit of the pleasure, though the tendency may be overpowered by an external or internal restraint" (1963, p. 215). Rather than treating valuing as a state that is conceptually tied to motivation as suggested in the text, we could see it (as Mill sees desire) as "the initiatory stage of the Will".

²⁰It is important to emphasize that this conceptual connection is merely *part* of the concept of intrinsic valuing. Taking it to be *definitive* of the concept of intrinsic valuing creates (or at least exacerbates) a problem for Humean theories that Warren Quinn puts as follows: showing that "pro- and con-attitudes conceived as functional states that dispose us to act have any power to rationalize those acts" (1993, p. 236). (Quinn's point is put in terms of dispositions to act rather than in terms of typical motivation, but his underlying concerns are independent of this difference.) This question of the *normativity* of Humean reasons is beyond the scope of this paper. However, there is no need to allow a possible misunderstanding that makes this problem seem harder to solve. Quinn seems correct in thinking that if pro-attitudes (here, the relevant ones are intrinsic valuing) are conceived of merely as a dispositions to act—or even merely as a disposition to be motivated to act—the problem of normativity appears intractable. As Quinn points out, a mere disposition to turn on radios in one's vicinity doesn't seem to give one "even a *prima facie* reason to turn on radios". But isn't it manifestly obvious that a *mere disposition* is not constitutive of having a pro-attitude toward that action (and, more particularly, that a mere disposition to bring about an end is not constitutive of intrinsically valuing that end)? A person's disposition to fall asleep at department meetings, get tongue-tied when talking with a celebrity, or display anger when talking to an ex-spouse hardly shows that the person intrinsically values these actions or any state of affairs to which she believes them to be instrumental. For all Quinn's examples show, the problem is not that certain pro-attitudes cannot rationalize actions but, rather, that these pro-attitudes cannot be adequately understood as merely dispositions to act or be motivated to act—an analysis that, to my knowledge, has not been proffered by any defender of a Humean position.

²¹Gibbard's insightful discussion (1990) of these matters suggests that our evolutionarily honed social nature may make us susceptible to the normative positions endorsed by others. This force, contingent on our sociality and the existence of a common "normative community", is of enormous

importance. It does not, though, exhaust our resources for rational appraisal and criticism. The charge of irrationality has, I believe, force beyond the limits of a normative community.

²² This point is inspired by Connie Rosati (1989), though she views the “self-invention” as a part of practical deliberation rather than a pre-condition for it. I think nothing of importance hinges on the terminology. What is at issue is how, if at all, the process is constrained by reason.

References

- Bentham, Jeremy. *Principles of Morals and Legislation*. (Darien, Connecticut: Hafner Publishing Co., 1970).
- Dancy, Jonathan. *Moral Reasons*. (Oxford: Blackwell Publishers, 1993).
- Darwall, Stephen. *Impartial Reason*. (Ithaca: Cornell University Press, 1983).
- Ewing, A.C. *The Morality of Punishment* (London: Kegan Paul, 1929).
- Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. (Oxford: Clarendon Press, 1990).
- Heath, Joseph. “Foundationalism and Practical Reason,” *Mind*, 106(1997)451-473.
- Hubin, Donald C. “Prudential Reasons,” *Canadian Journal of Philosophy*, 10(1980)63-81.
- . “Irrational Desires,” *Philosophical Studies*, 62(1991)23-44.
- . “Hypothetical Motivation,” *Noûs*, 30:1(1996)31-54.
- Hume, David. *A Treatise of Human Nature*. (Oxford: Clarendon Press, 1968).
- Kagan, Shelly. *The Limits of Morality*. (Oxford: Clarendon Press, 1989).
- . “On Instrumental Reasons”, presented at the Central Division Meetings of the American Philosophical Association, April, 1989.
- . *Normative Ethics*. (Boulder, CO: Westview Press, 1998).
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*, trans. by H. J. Paton. (New York: Harper & Row, 1956).
- Korsgaard, Christine. “Skepticism about Practical Reason”, *The Journal of Philosophy*, 83 (1986)5-25.
- Milgram, Elijah. “Was Hume a Humean?”, *Hume Studies*, 21(1995)75-93.
- Mill, John Stuart. *Miscellaneous Writings*. Ed. by John M. Robson. (Toronto: University of Toronto Press, 1963).
- Nozick, Robert. *The Nature of Rationality*. (Princeton, NJ: Princeton University Press, 1993).
- Quinn, Warren. *Morality and Action*. (New York, NY: Cambridge University Press, 1993).
- Rawls, John. “Two Concepts of Rules”, *The Philosophical Review*, 64 (1955).
- Rosati, Connie. *Self-Invention and the Good*. Ph.D. Dissertation, University of Michigan, (1989).
- Smith, M.B.E. “Is There a Prima Facie Obligation to Obey the Law”, *Yale Law Journal*, 82(1973)950-976.
- Vallentyne, Peter. “Teleology, Consequentialism, and the Past”, *The Journal of Value Inquiry*, 22(1988)89-101.
- Wittgenstein, Ludwig. *Philosophical Investigations*, 2nd edition. (New York: MacMillan Co., 1958).