

Hume's Internalism Reconsidered[†]

Dale Dorsey

Department of Philosophy
University of Kansas
1445 Jayhawk Boulevard
Wescoe Hall, rm. 3090
Lawrence, KS 66045
ddorsey@ku.edu

Final draft, forthcoming in the *Journal of Ethics and Social Philosophy*.

A practical reason is the sort of thing that is supposed to, as it were, “count in favor of” doing something. That a proposition p is true is reason to believe it. In the same way, the fact that some act is, say, morally required, prudentially required, aesthetically beautiful, etc., might be reasons to perform it. Intuitively speaking, if I could save millions from devastating poverty at a tiny cost to myself, I have a reason to do it—a reason that, again intuitively speaking, seems decisive. In this way it is proper to say that practical reasons are *normative*. Though morality, for instance, may require me to ϕ , I *ought* to ϕ if and only if I have sufficient practical reason to ϕ .

Though applying this set of conceptual categories to David Hume is certainly anachronistic, I want to inquire into Hume's understanding of practical reasons here. The standard reading of Hume's view holds, among other things, that he is a normative internalist; that, for Hume, legitimate practical reasons must be linked to an agent's set of desires or motivating passions.¹ Though “normative internalism” has a number of different incar-

[†]I am grateful for the kind assistance of John Bricke, Erin Frykholm, Elijah Millgram, Don Rutherford, Jenny Welchman, and two anonymous reviewers for JESP, whose comments were helpful and challenging.

¹This view has been advanced by Bernard Williams, “Internal and External Reasons” in *Moral Luck* (Cambridge: Cambridge University Press, 1981), though it's worth noting that Williams declares this account of practical reasons “sub-Humean”; David Sobel, “Subjective Accounts of Reasons for Action” in *Ethics* 111 (2001); John Rawls, *Lectures on the History of Moral Philosophy* (Cambridge, MA: Harvard University Press, 2000), 34; Christine Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 51-65; Houston Smit, “Internalism and the Origin of Rational Motivation” in *The Journal of Ethics* 7 (2003); Stephen Darwall, *British Moralists and the Internal 'Ought' 1640-1740* (Cambridge: Cambridge University Press, 1995); Donald Hubin “Desires, Whims and Values” in *The Journal of Ethics* 7 (2003); Nicholas Capaldi allows that Hume believed in normative obligations, but only insofar as they are derived from the

nations, normative internalism holds that one's practical reasons are—very roughly—determined by one's desire set. For normative internalism, a purported reason to ϕ is genuine if and only if the agent in question has a desire, the object of which ϕ helps to promote.² Because Hume is generally interpreted as holding that an agent's ends must be comprised of her desires, Humean internalism is equivalent to his instrumentalism, the view that one only has practical reason to do what promotes one's ends.³ Though the internalist, or instrumentalist, reading of Hume is popular, it gives rise to serious puzzles of interpretation. To pick one nearly at random, it appears that, on an internalist reading, Hume has serious difficulties establishing that the so-called “artificial” virtues of justice and promise-keeping are reason-giving, especially when it comes to characters like the sensible knave.⁴

I want to make trouble for this reading of Hume. In particular, I will show that the various passages (especially from the *Treatise*) that seem to indicate some form of normative internalism are inconclusive at best. Some, relying on these passages, have sought to show that Hume is, rather than an internalist, a “nihilist” or skeptic about practical reasons.⁵ Against the internalist and skeptical readings, I argue that there is substantial reason to believe that Hume's corpus is compatible with a more robust account of normativity than internalism allows. If so, we should be hesitant to suggest that Hume cannot solve various puzzles that arise on the assumption of some form of internalism. In particular, I will show that Hume has a genuine

agent's subjective motivations; see *Hume's Place in Moral Philosophy* (New York: Peter Lang Publishing, 1989), 116-7.

²For a more precise characterization, see Mark Schroeder's *Slaves of the Passions* (Oxford: Oxford University Press, 2008), chs. 2-3. It is worth noting that—though he is careful to not attribute his view to Hume himself—Schroeder calls his normative internalism the “Humean Theory of Reasons”.

³Like normative internalism, normative “instrumentalism” also means different things in different contexts. By “instrumentalism” here, I refer to the less innocuous thesis that one cannot have a practical reason to ϕ unless ϕ promotes one's ends, not the more innocuous thesis, labeled “pure instrumentalism” by Donald Hubin, that if one has a reason to promote the end, one has a reason to take the means to that end. Cf. Hubin, “What's Special about Humeanism” in *Noûs* 33 (1999). One could divorce internalism and instrumentalism if one denied the claim that one's ends must be made up of one's desires—a thesis central to Hume's project.

⁴See, for instance, David Gauthier, “Artificial Virtues and the Sensible Knave,” *Hume Studies* 18 (1992).

⁵Elijah Millgram, “Was Hume a Humean?” and “Hume, Political Noncognitivism, and *The History of England*” in *Ethics Done Right* (Cambridge: Cambridge University Press, 2005). See also Jean Hampton, “Does Hume have an Instrumental Conception of Practical Reason?” in *Hume Studies* 21 (1995); and Donald Hubin, “What's Special about Humeanism”, op. cit.

response to the sensible knave which establishes the knave's obligation to, among other things, justice.

The organization of this paper runs as follows. In §1, I discuss the various reasons one might believe that Hume subscribes to some form of normative internalism. In §2, I show how these passages support normative internalism only on dubious interpretive assumptions. Next, I attempt to show that Hume subscribes to two crucial features of the denial of normative internalism. First, Hume attempts to criticize desires and motivations from a point of view that is fully independent of the agent's own subjective motivational set. Second, he believes that this point of view establishes normative obligations on the part of agents. In §4, I show how this interpretation can establish a plausible reading of Hume's treatment of the sensible knave in the final section of the *Enquiry*. In §5, I discuss the charity of the resulting view, and in §6, I conclude with some very brief thoughts concerning the contemporary relevance of Hume's considered position.

1. *Hume's Internalism*

Before diving into various passages that seem to point to Hume's normative internalism, it is worthwhile to say a bit about what I mean, and what various interpreters of Hume mean, when referring to "normative internalism." Normative internalism comes in many shapes and sizes, some weaker, some stronger. The tie that binds is a necessary connection to motivation: a purported reason for x to ϕ must maintain a necessary connection to a state of motivation. But stated in this way, normative internalism is compatible with every possible account of practical reason. For instance, the following view—which is compatible with all views—maintains a necessary connection to motivation: a purported reason for x to ϕ is a genuine reason for x to ϕ if and only if ϕ would motivate an agent who is motivated by genuine practical reasons.⁶

But in considering *Hume's* purported normative internalism, I certainly don't mean anything this weak. Though it is uncontroversial that Hume does not suggest that every possible motivation provides genuine practical reasons, genuine practical reasons must find their home in the preexisting subjective motivational or desiderative set of the agent. Certain motivations will fail to provide reasons, i.e., motivations that are based on false beliefs,

⁶Though Korsgaard does not embrace a view that is *quite* this weak, her account of internalism in "Skepticism about Practical Reason" (in *The Journal of Philosophy* 83 (1986), esp. 22) is quite similar.

or that do not serve our ultimate ends, or serve them inefficiently (T 2.3.3.6). But in claiming that Hume is a normative internalist, normative internalism is generally understood to mean something like the following: a purported reason for x to ϕ is a genuine reason for x to ϕ if and only if ϕ promotes some ultimate (non-derived) part of x 's motivational set, and is not based on false beliefs or incorrect assessments of the most efficient means to our ends. (For short, I will simply say "false beliefs".)

One further way of bringing the standard picture of Hume's internalism into focus is to consider, on this version of internalism, the extent to which motivations themselves are subject to criticism. On the view proposed in the previous paragraph, any rational criticism of desires must make essential reference to *other* desires possessed by the agent in question. One might criticize, for instance, a desire that was formed on the basis of some false beliefs or other, and is therefore in conflict with a primary or ultimate motivations. Alternatively, one might criticize a motivation for being too costly with regard to one's overall desire set. For instance, one might desire to eat a candy bar, but might have a stronger desire to avoid hyperglycemia. Or one might have such a range of desires that require one to forsake the candy bar that eating the candy bar—in terms of one's overall motivational set—just isn't worth it. In this case, the desire to eat the candy bar is normatively *outweighed* by other, stronger, motivations. One helpful way of putting this point is by making a distinction between *pro tanto* reasons and *all things considered* reasons; all underived desires provide *pro tanto* reasons to promote their objects; not all underived desires provide *all things considered* reasons to promote their objects. Sometimes the rest of one's desire set might simply outweigh the normative importance of acting on a particular desire. Further, one might have reason to ignore or eliminate a desire, or to develop a desire if adding or subtracting the desire in question would have beneficial effects when it comes to the overall fulfillment of the agent's desire set.⁷ Though this range of criticism is available to the normative internalist, the internalist must admit that all *pro tanto* practical reasons are derived from an agent's subjective motivational set. One denies normative internalism if and only if there are at least some *pro tanto* practical reasons that do not require a corresponding desire. If one allows that a desire can be rationally criticized without making essential reference to other facets of the subjective motivational set of the agent in question, one denies normative internalism.

⁷Cf. Hubin, "Groundless Normativity of Instrumental Rationality" in *The Journal of Philosophy* 98 (2001), 456-7.

Though there is much more that could be said about various versions and refinements of normative internalism, it seems to me that something like this view is generally held up as the standard picture of Hume’s theory of practical reasons. For instance, John Rawls writes: “Thus reasons for action must connect, it seems, with one or more of our *existing* passions. This is one thing meant by speaking of Hume’s view as *internalist*: what count as reasons for someone must link up with *that* person’s currently existing motivations, in Hume’s case, with that person’s currently existing passions.”⁸ David Sobel writes: “Acts, for Hume, are contrary to reason, albeit indirectly and in an ‘improper’ way of speaking, when the act is motivated by a passion which would not exist (or, presumably, would be significantly altered) except for misinformation. On this reading of Hume it would be natural for him to say that one acts in accord with reason or one has a genuine reason to ϕ when one’s motivation to ϕ is not based on some such misinformation.”⁹

With this in mind, there are two central pieces of evidence, one from Hume’s *Treatise*, and one from the *Enquiry*, that seem to point in the direction of some form of normative internalism. The first set of citations comes from Hume’s discussion in “Of the Influencing Motives of the Will” (T 2.3.3). In a headline-grabbing paragraph, Hume appears explicitly to deny that any form of criticism is applicable to passions beyond criticism that stems from these passions being based on false beliefs. Hume writes, famously, that

’tis only in two senses, that any affection can be call’d unreasonable. *First*, When a passion, such as hope or fear, grief or joy, despair or security, is founded on the supposition of the existence of objects, which really do not exist. *Secondly*, When in exerting any passion in action, we choose means insufficient for the design’d end, and deceive ourselves in our judgment of causes and effects. Where a passion is neither founded on false suppositions, nor chooses means insufficient for the end, the understanding can neither justify nor condemn it. (T 2.3.3.6)

Recall that one feature of internalism was the reluctance of internalists to subject motivations to critical scrutiny external to the desire set itself (which will often consist in noting failures of theoretic rationality, failures of means-ends reasoning and/or false beliefs). Here Hume appears to explicitly rule

⁸Rawls, op. cit.

⁹Sobel, 468.

out additional errors when it comes to motivations or passions. If these errors are not possible, Hume is cast in the role of a normative internalist; internalism's denial requires that motivating passions can be criticized to from a standpoint external to an agent's desire or motivational set. But here, as Hume says, "a passion must be accompany'd with some false judgment, in order to its being unreasonable; and even then 'tis not the passion, properly speaking, which is unreasonable, but the judgment."

The above passages seem to indicate that Hume refused to countenance criticism of desires based on some desire-independent standard, whether that be the standard of practical reason or some other standard. But there is further evidence that suggests that even *if* Hume allowed such criticism, he would not have regarded that criticism as normatively important. In other words, even if passions could be criticized, the passions themselves would continue to be the sole source of *practical reasons*.

The second major passage is found in the first Appendix to the *Enquiry*, Hume writes:

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 will readily reply, *because sickness is painful*. If you push your enquiries farther, and desire a reason, *why he hates pain*, it is impossible he can ever give any. This is an ultimate end, and is never referred to any other object.

Perhaps, to your second question, *why he desires health*, he may also reply, that *it is necessary for the exercise of his calling*. If you ask, *why he is so 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. It is impossible there can be a progress *in infinitum*; and that one thing can always be a reason, why another is desired. Something must be desirable on its own account, and because of its immediate accord or agreement with human sentiment and affection. (EPM App 1.18-19.)

Here, as in the passages from the *Treatise*, Hume appears to explicitly rule out applying "reasons" to one's ultimate-non-derived-desires. As far as reasons are concerned, once one has hit the ultimate desire, one has hit rock bottom; there is no other sense in which a reason can be offered for a desire except that that desire is in accord with the agent's *other* desires.

Though these passages seem to point directly at normative internalism, I argue that they are, at best, inconclusive. I argue that Hume's passages

here point to a normative criterion that is significantly more robust than the form of internalism often ascribed to Hume.

2. Reason, Reasons, and “Reasons”

I take the above passages in order. In particular, the most influential passage in establishing Hume’s normative internalism has been the discussion at, and surrounding, T 2.3.3.6.¹⁰ But the discussion in which Hume frames these comments is not a discussion that is directed toward the concept of practical reason, *per se*. Hume’s goals are, of course, to show that “reason” can never influence the will—that passions themselves cannot conflict with “reason” (thus establishing that there is no combat between reason and passion). Now, this would support normative internalism if and only if Hume’s “reason” here *included* all normative assessment of passions and actions. But we should not believe this. In particular, at T 2.3.3.2, Hume writes: “The understanding exerts itself in two different ways, as it judges from demonstration or probability; as it regards the abstract relations of our ideas, or those relations of objects, of which experience only gives us information.” Later, Hume writes that “Reason or science is nothing but the comparing of ideas, and the discovery of their relations,” (T 3.1.1.24). Here, according to Hume, reason simply *is* science—the empirical discovery of facts about our world.

What does this suggest about the headline-grabbing passages at T 2.3.3.6? Importantly, it suggests that *nothing in Hume’s text implies that passions are normatively authoritative merely because they cannot be unreasonable*. If passions can be described as unreasonable only when those passions are based on false beliefs, this entails only that *certain* forms of criticism fail. In particular, all this establishes is that nothing about empirically discoverable facts about the nature of our world could be used to criticize our passions. Passions cannot be criticized by *reason*. But one could suggest that passions can be criticizable in a further sense, even if they can never, strictly speaking, be unreasonable. For instance, one could suggest that passions themselves display not errors of reason, but errors of *passion*. Consider two agents. One prefers Pabst Blue Ribbon to Guinness, the other prefers Guinness to PBR. Nothing in what Hume suggests at T 2.3.3 entails that we could not criticize these various passions, or that there is nothing wrong

¹⁰Those who take these passages to be conclusive include Sobel, Rawls, and David Brink, “The Significance of Desire” forthcoming in *Oxford Studies in Metaethics*, ed. Shafer-Landau (Oxford: Oxford University Press, 2008). Brink interprets Hume as a skeptic.

with preferring PBR to Guinness or vice versa.

Misunderstanding might result from Hume's terminology. Hume's use of the term "reason" is as a *faculty*—the faculty of reason, by which we ascertain various facts about the world. Reason, construed as a faculty, cannot be practical on Hume's view.¹¹ For others, including Kant, the capacity or process of mind is not limited—it can be practical as well as theoretical. In this sense, Hume surely *is* a skeptic about practical reasoning. But merely because Hume rejects Kantian-style rationalism does not mean that he cannot accept that independent practical *reasons* exist, construed as factors that count in favor of doing something or other. In particular, considering the example of the previous paragraph, one might insist that the mistaken passions of the PBR fan are reasons to develop new passions, or to avoid loud praise of PBR over Guinness in polite company.¹²

A confusion here can be spotted in Millgram's treatment of Hume's passages at T 2.3.3.6. Millgram seeks to argue that, for Hume, reason—construed as a faculty—is never practical. But on Millgram's view, Hume's account amounts to the "null hypothesis" that actions cannot be "rational" or "irrational". Millgram writes:

One is either extremely fortunate or unfortunately complacent if one has not had bleak mornings during which it seems suddenly clear that purported reasoning about action is nothing more than empty posturing, the attempt to proceed under the comforting but unsupported notion that actions or decisions, or the mental activities leading up to them, might be right or wrong, because rational or irrational.¹³

It is true that mental activities cannot be right or wrong, for Hume, because they fail to conform to the faculty of reason. Furthermore, "purported rea-

¹¹Annette Baier has argued that Hume holds that the faculty of reason can be imbued with the responses from our passions and sentiments, and will for this reason be a reliable guide to, among other things, morality. But at this point in the *Treatise*, it is reasonably clear that Hume intends reason to mean, as Baier puts it, a faculty of mind that includes only "causal reasoning and probability estimation." See *A Progress of Sentiments* (Cambridge, MA: Harvard University Press, 1991), 277-8.

¹²One natural way to mark the difference between these two uses of "reason" is linguistically: "reason" in the sense of a capacity or process of mind does not permit of a plural form. "Reason" in the sense of "that which counts in favor of ϕ -ing" does permit of a plural form. I can have multiple, even conflicting, "reasons" (i.e., I can have a reason to ϕ and a reason to ψ , where ψ entails the failure to ϕ). I don't have a capacity of "reasons", but rather a capacity of "reason".

¹³Millgram, "Was Hume a Humean?", 212.

soning about action,” *can*—in a sense—be properly described as “empty posturing” so long as “reasoning” is essentially linked to the faculty of reason. In this, Millgram is correct. But this does not imply that Millgram’s “null hypothesis” was held by Hume. For all that has been said so far, Hume *could* believe that actions can be right or wrong because they fail to conform to practical *reasons*. In this sense, passions can be “irrational”—they can fail to live up to the practical reasons we have. These practical reasons need not be derived from the faculty of *reason*. Insofar as those who would interpret Hume as a skeptic about practical reason confine their discussion to reason *qua* faculty, I have no beef. But it must be understood that this goes no distance whatever in establishing whether Hume was a *normative* skeptic, a skeptic about practical *reasons*. In making this point, I agree with David Phillips, who writes, concerning the passages at 2.3.3, “That passions cannot represent, and hence cannot be true or false, does not entail that they cannot be correct or incorrect. It entails only that their correctness or incorrectness can’t be a matter of truth or falsity, of representational adequacy or inadequacy. It does not entail that there are no authoritative norms applying to passion or action.”¹⁴

Others have noticed the shift of language between Hume’s use of the term “reason” and our contemporary understanding of “practical reasons”. In discussing a suggestion by Parfit, Kieran Setiya writes:

[I]n Book Two, and in particular, in the section “Of the influencing motives of the will”, Hume is using “reason” as he does at the beginning of Book Three. He means the capacity for theoretical or truth-directed reasoning: that is, on his particular account of it, the capacity for demonstrative and causal inference. So, in arguing that passions cannot be contrary to reason, Hume does not illicitly assume “that there is only one kind of reason: reasons for believing”.¹⁵

The quotation from Parfit commits precisely the fallacy I note here. Hume is not discussing “reasons” *for* anything in T 2.3.3.6. Rather, he is discussing the extent to which the faculty of *reason* might influence the will. Hume’s answer is that reason cannot influence the will, because passions are not properly spoken of as compatible or incompatible with the verdicts of science

¹⁴David Phillips, “Hume on Practical Reason: Normativity and Psychology in *Treatise* 2.3.3” in *Hume Studies* 31 (2005), 305. See also Kieran Setiya, “Hume on Practical Reason” in *Philosophical Perspectives* 18 (2004), 370.

¹⁵Setiya, 370; see also 373. The quotation is from Derek Parfit, “Reasons and Motivation” in *Proceedings of the Aristotelian Society, Supplementary Volume* 71 (1997), 128.

and mathematics. But this does not entail that an individual's motivational set is normatively sacrosanct, that desire can only be criticized on the basis of other desires. Setiya is correct, and his interpretation fits Hume's text better than the internalist, or skeptical, alternative. Thus, it seems to me, we should be hesitant to conclude that Hume's discussion at T 2.3.3 supports normative internalism. As I shall suggest in the next section, Hume himself held that there were good reasons for criticizing passions beyond their merely being in conflict with other desires or motivational passions.

What, then, to say concerning the problematic passage at EPM App 1.18-19? Hume appears to be using "reason" here in a way that only awkwardly applies to a faculty. Beyond some particular sentiment or desire, says Hume, it is "an absurdity to ask for a reason." He appears not to be referring to reason *qua* faculty here, but rather reasons one might hold a particular desire or other.

In response, it is helpful to note that ordinary English, at which I have already gestured, has many senses of "reason". If I ask for a "reason" *for* a desire, I could be asking for a *normative* reason, that is, whether I should have that desire, or what justifies me in having such a desire. Alternatively, I might be asking for an *explanation* of the desire. Just as I might explain the reason *for* the low temperatures (low pressure system, winter, northern latitude, etc.), I might explain the reason *for* a desire by citing some causal/explanatory story. And it is the causal/explanatory sense that Hume means in this passage. The passage in question seeks to establish that "human actions can never, in any case, be accounted for by *reason*, but recommend themselves entirely to the sentiments and affections of mankind, without any dependence on the intellectual faculties," (EPM App 1.18). Here Hume is seeking to *account* for or *explain* human actions by noting that the only thing that could do so is the agents' desires. Hume here is seeking an account, an explanation, first of action, and then of the desires that give rise to a particular action. The function of the series of "why?" questions Hume asks and answers is explanatory, not justificatory—why it is that this particular agent desires *x* rather than *y*. This passage succeeds, to my mind, in establishing not normative internalism, but rather the "Humean Theory of Motivation", viz., that desires are required in order to explain motivated action.

But even if one wanted to draw a specifically normative conclusion from Hume's discussion at EMP App. 1.18-19, the normative conclusion Hume wishes us to draw is far weaker than is required for normative internalism. Hume writes that actions "recommend themselves entirely to the sentiments and affections of mankind, without any dependence on the intellectual fac-

ulties.” Further, he says: “Something must be desirable on its own account, and because of its immediate accord or agreement with human sentiment and affection.” Hume does *not* say that actions are “desirable” for a particular agent only insofar as they engage the motivational set of that *particular* agent. Rather, his claim is more general: ϕ ’s desirability is to be established by ϕ ’s accord with “human sentiment and affection”. And, as I shall show in the following section, Hume argues that “human sentiment and affection” can be an effective means of critically engaging the motivational passions of particular agents.

Thus there is still room for Hume to believe that passions and desires can undergo external normative criticism. In the passages cited from the *Treatise*, Hume seeks to establish that *reason* is not practical. As it happens, Hume seeks to establish the same conclusion in the above passage from the *Enquiry*’s first Appendix. Rather than arguing that desires cannot be externally criticized, Hume argues that ends can only be explained by reference to the ultimate desires of agents. However, I wish to argue in the following section that Hume clearly *did* believe that our desires are subject to external normative criticism. For Hume, we can have practical reasons that are not dependent upon our antecedent motivations.

3. *The Normativity of Taste*

This section will answer two questions. The first question, suggested by my discussion of T 2.3.3 above, concerns whether Hume allowed that motivations are subject to independent sources of criticism. I think the answer to this question is yes. However, a “yes” answer to the first question is insufficient to establish that Hume is not a normative internalist. One also has to show that the standard of criticism to which one’s motivations are subject establishes reasons to *change* one’s motivations or behavior. To put this another way, one must establish that the standard against which one’s own passions can be criticized is genuinely reason-giving. Without this second project, whether Hume is a normative internalist remains up for grabs.

3.1. *Hume’s Aesthetics and the Criticism of Passion*

The point at which Hume most clearly criticizes the passionate responses of agents is in his discussion of aesthetic criticism in “Of the Standard of Taste”. Hume is keen to dispute the view that holds that “All sentiment is right; because sentiment has a reference to nothing beyond itself, and is always real, wherever a man is conscious of it,” (OST 230). Hume’s intu-

itions are on the side of those who declare that “Whoever would assert an equality of genius and elegance between Ogilby and Milton, or Bunyan and Addison, would be thought to defend no less an extravagance, than if he had maintained a mole-hill to be as high as Teneriffe, or a pond as extensive as the ocean,” (OST 230-1). But if that is the case, Hume has got to tell us the means by which the passionate responses of those who *do* so assert this equality are mistaken or untrustworthy. And this he seeks to do: “It is natural for us to seek a *Standard of Taste*; a rule, by which the various sentiments of men may be reconciled; at least, a decision, afforded, confirming one sentiment, and condemning another,” (OST 229).

Though this essay is wonderfully complex, and permits of many problems of interpretation, it is at least relatively clear that Hume believes that those whose sentimental reactions fail the standard of taste have sentiments that can be criticized *by* that standard. Those whose reactions depart from the standard of taste, according to Hume, lose “all credit and authority,” (OST 240). According to Hume, this standard is made up of several conditions, not all of which will be possessed by all humans at all times, some of which require intense and significant regimens of cultivation. Hume spends the most time discussing the so-called “delicacy” of taste, the ability to make fine distinctions. However, Hume also notes that persons will fail the standard of taste when they are under the spell of some form of prejudice, either for themselves or for the time or place in which they live. Additionally, Hume insists that the proper standard of taste requires sufficient practice and proper comparison between things of purported beauty. “Strong sense, united to delicate sentiment, improved by practice, perfected by comparison, and cleared of all prejudice, can alone entitle critics to this valuable character; and the joint verdict of such, wherever they are to be found, is the true standard of taste and beauty,” (OST 241).

It is clear, then, that Hume believes that the passions can be criticized for failing to conform to the verdicts of the standard of taste. He goes so far as to suggest that these departures are “blameable” (OST 246, 247), and “condemnable,” (OST 229). But it might be noticed that Hume’s topic here is aesthetics and art criticism. Does Hume’s discussion of aesthetics entail criticism of our *motivational* sentiments or desires?

This point is key. One might make a distinction in the types of sentiments we have; critical sentiments allow us to distinguish beautiful from ugly things and properly identify them as such. Motivational sentiments, on the other hand, are those things that give rise to action. If Hume is willing to grant that our assessment of moral characters is subject to the standard of taste, but is unwilling to grant that the motivations of those characters assessed

are similarly criticizable, it would be difficult to establish the claim that Hume is anything but a normative internalist; motivations would not be subject to external normative assessment.

But this move is available only if Hume marks a distinction between motivational and critical sentiments. He does not. In particular, Hume notes at EPM App 1.21, that “taste...gives the sentiment of beauty and deformity, vice and virtue.” But a few lines later, Hume declares that the same set of sentiments that mark moral distinctions are also motivational: “Taste, as it gives pleasure or pain, and thereby constitutes happiness or misery, becomes a motive to action, and is the first spring or impulse to desire and volition.” Hume, it appears, refrains from marking any philosophical or psychological distinction between “actuating” passions—those that lead to action—and critical passions—those that provide a “sentiment of beauty and deformity”. On Hume’s view, there is a standard of *taste simpliciter*—both our “motivational” sentiments and “critical” sentiments fall under this heading. Later in that same paragraph, Hume discusses the standard of taste (which, on Hume’s view, arises “from the internal frame and constitution of animals”)—a standard that applies both to critical sentiments as well as sentiments that engage us to act. Thus it appears that the standard of taste applies not only to our critical sentiments, but also to our motivational sentiments. Just as I might be justly criticized for praising a bank robber as instantiating moral beauty, I might be criticized for desiring to rob a bank, or desiring to act in a way that is in some other way morally ugly.

The argument of this section was to some degree complex, so it might help to state it explicitly here.

1. Hume subjects our aesthetic and critical sentiments to an independent standard. (OST.)
2. But Hume insists that our moral sentiments function in the same way as our aesthetic sentiments. (EPM 1.9, EPM App 1.13-16.)
3. Hence, Hume subjects our moral sentiments to an independent standard. (From 1 and 2.)
4. But Hume makes no distinction between our sentiments that assess character, i.e., our moral sentiments, and our motivational sentiments. (EPM App 1.21.)
5. Hence, Hume subjects our motivational sentiments to critical assessment by an independent standard. (From 3 and 4.)

3.2. *Taste and Reasons*

Though Hume subjects our motivations to independent, critical examination *via* the standard of taste, we still have not yet established that Hume is not a normative internalist. The crucial question is this: is the standard of taste reason-giving? If ϕ would motivate under conditions conducive to the standard of taste, do I—at least occasionally—have reason to ϕ , despite my own (perhaps opposed) motivations? In this section, I wish to argue that Hume answers in the affirmative.

The first point to make here is that Hume explicitly ties his account of aesthetics and his account of the moral sentiments together in the *Enquiry*. In EPM 1.9, Hume asserts that “in many orders of beauty, particularly those of the finer arts, it is requisite to employ much reasoning, in order to feel the proper sentiment; and a false relish may frequently be corrected by argument and reflection. There are just grounds to conclude that moral beauty partakes much of this latter species, and demands the assistance of our intellectual faculties, in order to give it a suitable influence on the human mind”. (See also EPM App 1.13-16.) Of course, this is but one (defeasible) similarity. Later, however, Hume explicitly links the standard of taste with *both* aesthetic judgments and moral judgments. Taste, writes Hume, provides us with “the sentiment of beauty and deformity, vice and virtue,” (EMP App 1.21). Hence if *taste* gives rise to sentiments in both domains, it would appear that the *standard* of taste applies not only to aesthetics, but also to morality. Furthermore, Hume explicitly embraces certain “corrections” of moral sentiment that mirror his discussion in the standard of taste. To note one example, Hume insists on tempering the extent to which moral judgments betray a prejudice to one’s own time, place, or interests (EPM 5.42). There is good reason, then, for concluding that judgments of *moral* beauty, according to Hume, are subject to precisely the same standards of taste as judgments of *aesthetic* beauty. On Hume’s view, moral judgments must be properly practiced, free of prejudice, competent in making comparisons, and sufficiently delicate in order to arrive at the authoritative verdict. Authoritative moral sentiments are a *subset* of the verdicts of the standard of taste.

Thus if Hume is to deny that the standard of taste provides reasons for action in cases in which an agent’s sentiments diverge from the standard, he must subscribe to the view, call this “normative skepticism”, that morality can sometimes fail to provide practical reasons, can fail to be normative. Given that Hume, as noted above, explicitly suggests that moral sentiments are subject to verification by the standard of taste, he would have to believe that morality can fail to be normative if the standard of taste fails to be normative, i.e., in cases in which an agent’s motivational set is not advanced

by conforming to the verdicts of the standard. And perhaps there is evidence for the claim that Hume subscribed to this form of normative skepticism about morality. At least in the *Treatise*, Hume appeared to be substantially gun shy concerning whether his inquiry into moral theory was intended to be practical. Michael Gill suggests that Hume cared very little about instilling virtue in his readers; his task was to describe a system of norms, not to establish that this system provides reasons for action. Gill writes: “The *Treatise* differs from the writings of earlier British moralists, first of all, in being an essentially *theoretical* work, not a practical one. The goal of the *Treatise* is to provide an account that best captures the observable phenomena of human behavior. It does not try to convince people that they ought to act in certain ways.”¹⁶ It is always open to us, of course, to suggest that Hume changed his mind. Further, it is open to us to believe that Hume intended to treat the *Treatise* as an apractical examination of a normative subject. But, if Gill is correct, Hume’s refusal to treat the *Treatise* as practical might give us pause when it comes to interpreting Hume’s normative stance toward morality, especially if, as Gill writes, the *Treatise* is unique among works of the British Moralists for its refusal to treat morality as practical.

But I think Gill is too quick in suggesting that Hume didn’t intend his moral inquiry to have practical implications. In particular, at 3.3.6.6, Hume makes an analogy between an anatomist and a painter. “The anatomist ought never to emulate the painter; nor in his accurate dissections and portraitures of the smaller parts of the human body, pretend to give his figures any graceful and engaging attitude or expression. There is even something hideous, or at least minute in the views of things, which he presents; and ’tis necessary the objects shou’d be set more at a distance and be more cover’d up from sight, to make them engaging to the eye and imagination.” Here Hume *explicitly* rejects the suggestion that he is trying to make morality look engaging to the reader; his task is only that of the anatomist, spilling our moral and psychological guts out for all to see. However, in elaborating on the analogy, Hume insists that the moralist (painter) has much to learn from the anatomist. On Hume’s view, the anatomist provides essential information to those who would make beautiful and engaging illustrations of the body.¹⁷ Finally, Hume declares that “the most abstract speculations concerning human nature,” of the kind in which Hume has been engaging,

¹⁶Michael Gill, *The British Moralists on Human Nature and the Birth of Secular Ethics* (Cambridge: Cambridge University Press, 2006), 201.

¹⁷Cf. Robert Shaver, “Hume’s Moral Theory?” in *History of Philosophy Quarterly* 21 (1995), 319-20.

“however cold and unentertaining, become subservient to *practical morality*; and may render this latter science more correct in its precepts, and more perswasive in its exhortations.” Here it is clear that Hume believed that his *Treatise* was at least indirectly practical—he believed that though he himself was no moralist, no moralist could succeed without learning from the conclusions of his *Treatise*. Thus Hume did not believe that the *Treatise* is devoid of practical application. However, as noted before, even if we agree with Gill, this says little about normative skepticism when it comes to morality. *Morality* can be perfectly practical even if the *Treatise* is not; the true account of morality (i.e., the *Treatise*) simply needs to be entrusted to more persuasive artists.

Leaving aside the *Treatise*, however, it is clear that he intends “moral beauty” to be reason-giving once we get to the *Enquiry*. At EPM 1.7, Hume specifically declares that true moral claims seek to teach us our normative obligations, i.e., how we should live. “The end of all moral speculations is to teach us our duty; and, by proper representations of the deformity of vice and beauty of virtue, beget correspondent habits, and engage us to avoid the one, and embrace the other.” There should be very little doubt, then, that Hume regards the requirements of morality as practical: morality’s requirements are to be lived up to. Morality, for Hume, teaches us our duty and engages us to actually perform it—it is practical in at least this sense. Were Hume to believe that morality was not practical, or could fail to provide practical reasons, recommendations to conform to moral requirements or to “beget corresponding habits” would seem perverse.

Furthermore, the suggestion that Hume regards the standard of taste as normative passes the test for normative obligations that Hume sets in the *Treatise*, especially T 3.2.5.6. Famously, Hume writes, “No action can be requir’d of us as our duty, unless there be implanted in human nature some actuating passion or motive, capable of producing the action.” Further, at T 3.2.1.7, he writes: “In short, it may be establish’d as an undoubted maxim, *that no action can be virtuous, or morally good, unless there be in human nature some motive to produce it, distinct from the sense of its morality.*” Hume suggests that obligations require a motivation found in human nature. But it is absolutely crucial to note that, for Hume, the standard of taste *is* found in human nature. Though agents may sometimes depart from the standard of taste, this standard is rooted in the common, shared features of humankind. On this point Hume is quite clear: whatever characterizes the proper standard of taste, that standard is rooted in a conception of human nature: “The general principles of taste are uniform in human nature: Where men vary in their judgments, some defect or perversion in the faculties may

commonly be remarked; proceeding either from prejudice, from want of practice, or want of delicacy; and there is just reason for approving one taste, and condemning the other,” (OST 243; see also EPM App 1.21). This claim is important, and worth underlining, for a further reason. If the standard of taste is nothing more than the considered opinion of, say, some statistical majority, it is certainly reasonable to wonder why it might provide practical reasons at all. Put another way, if the standard of taste is just some group’s view, it might be thought manifestly implausible for that standard to be normative for all. But this is not Hume’s view. As Hume makes clear, the standard of taste finds its home within human nature itself—departures from the standard are considered defects. In suggesting that the standard of taste provides practical reasons, Hume is arguing that we have reason to eliminate these defects, to develop and act in accordance with our basic human nature. If so, it is far less implausible to believe that the standard of taste is normative.¹⁸

One might then ask about whether the account of “normativity” here is anything we’d really recognize as an account of practical reasons. After all, shouldn’t practical reasons be the sort of thing that are accessible to agents? What *use* is it to have practical reasons applying to agents that aren’t motivated to act on them? I think the answer is that it is not at all puzzling to ascribe reasons for action to agents who do not/could not act on them. Rationalists, no less than sentimentalists, insist on the existence of practical reasons that do not motivate action. Imagine that some agent has the opportunity to save the entire continent of Africa from horrible starvation if only he would give up his lollypop. A racist and a lover of lollypops, this person refuses, and could not be convinced. In this case, we would like to say, the agent has an overwhelming practical reason to give up the lollypop, though he could not be convinced or motivated to act otherwise. On a rationalist account, this person is not motivated because the person fails rationality—this person either fails to possess the capacity for rational agency or deliberation, or is *akratic* in some way or other. However, this point is also plausibly explained on Hume’s view: the desire to give up the lollypop, though not possessed by this agent, is a desiderative verdict of the standard of taste, and can hence support a genuine reason for this agent to act. Though this person may possess a fully developed capacity for rationality, he does not possess the crucial moral capacity: a developed taste. Hume’s view on practical reasons and their accessibility should not

¹⁸Thanks to an anonymous reviewer for urging me to discuss this issue in more detail.

be seen as mysterious or at all uncharitable.¹⁹

In addition, this account of Hume’s view is able to plausibly accommodate a test that Korsgaard attributes to Hume. As Korsgaard states it, “a faculty’s verdicts are normative if the faculty meets the following test: *when the faculty takes itself and its own operations for its object, it gives a positive verdict.*”²⁰ A similar view is attributed to Hume by Baier.²¹ But given Hume’s identification of moral beauty with the verdicts of the delicate, non-prejudiced, practiced, taste, moral principles are normative because the delicate, etc., sentiment approves of its operations *vis-a-vis* moral principles. Thus for Hume, desires can be rationally criticized without making essential reference to the agent’s subjective motivational set. Desires can be rationally criticized for not conforming to the verdicts of taste. This view reflects Hume’s texts better than the internalist competitor.²²

It is worth spelling out the argument of this section in more detail. Very briefly, I have argued:

1. For Hume, the conditions of sentimental authority in aesthetics are identical in the conditions of sentimental authority in morality. (EPM 1.9, EPM App 1.13-16, 21.)
2. Hence, Hume believes that moral sentiments are authoritative only when possessed of the standard of taste. (From 1 and OST.)
3. Hence, if Hume believes that no verdicts of the standard of taste are normative for agents with contrary motivations, Hume is a normative skeptic about morality. (From 2.)
4. But Hume is not a normative skeptic about morality. (T 3.3.6.6, EPM 1.7.)
5. Hence, Hume believes that at least some verdicts of the stan-

¹⁹Thanks to Elijah Millgram for pressing me on this point.

²⁰*The Sources of Normativity*, 62.

²¹*A Progress of Sentiments* (Cambridge, MA: Harvard University Press, 1991), 97-101.

²²An internalist might respond that the standard of taste could be normative for individuals who are not so motivated if those individuals, for instance, have a *de dicto* motivation: to behave in accordance with good taste. But it is worth noting that, though many people will have this *de dicto* motivation, Hume’s account of the normativity of morality does not appear to depend on *any* facts about the subjective motivational sets of agent; Hume never subjects the reason-giving force of morality to a *de dicto* motivational caveat. There may be many ways that the internalist can accommodate a reason to conform to the standard of taste—but none can do so independently of the subject’s desires or motivations. Further evidence that Hume divorces practical reasons and motivations—even *de dicto* motivations—is presented in §4, in the form of the sensible knave. Thanks to an anonymous reviewer for pointing out this possibility.

- dard of taste are normative. (From 3 and 4.)
6. But the verdicts of the standard of taste do not motivate all agents. (From OST.)
7. Hence, Hume believes that not all practical reasons must be motivating. (From 5 and 6.)

4. *Prudence, Reasons, and the Sensible Knave*

My account of Hume’s normative resources does work. In particular, it appears to solve an impasse between the demands of morality and the sensible knave that has troubled interpreters of Hume. Hume appears to admit, at EPM 9.2.22 that self-interest is not sufficient to establish a reason for agents to conform their actions to justice. According to Hume, it is possible that “taking things in a certain light,” an agent “may often seem to be a loser by his integrity.” And though honesty and fidelity to property are often important in establishing a working society, upon which everyone’s interests depend, we could imagine a sensible knave whose cunning is sufficient to allow him to free ride, perhaps undetectedly. For this person, “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.”

One obvious question to ask is, if the sensible knave’s motivations are organized as such, is the sensible knave obligated to behave in accordance with justice? Hume appears to be at something of a loss to respond to the sensible knave, or to characters who might be tempted to defect from the moral cause: “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 will to him appear satisfactory and convincing. If his heart rebel not against such pernicious maxims ... we may expect that his practice will be answerable to his speculation,” (EPM 9.2.22). On this point, one might compare Hume’s discussion of a knavish character in “The Sceptic”: “On the other hand, where one is born of so perverse a frame of mind, of so callous and insensible a disposition, as to have no relish for virtue and humanity, no sympathy with his fellow-creatures, no desire of esteem and applause; such a one must be allowed entirely incurable, nor is there any remedy in philosophy... For my part, I know not how I should address myself to such a one, or by what arguments I should endeavour to reform him.”²³

If this is Hume’s answer, a wealth of questions pile up. The most obvious question is the normative question. Does the sensible knave have a *reason* to behave in accordance with justice? Responding as Hume does

²³ “The Sceptic” in *Essays: Moral, Political, and Literary*, op. cit., 169.

above seems to flirt with a “no” answer: if we can “expect that his practice will be answerable to his speculation,” it seems unlikely that he will be motivated to act justly. But this is especially puzzling given that, in EPM 9.2, Hume is attempting to establish that we have an “interested obligation” to justice. In other words, Hume seeks to answer the normative question in the affirmative: not only does the sensible knave have a reason to behave justly, he has a *prudential* reason to behave justly. Hume seeks to establish that his account of morality “can show, by particular detail, that all the duties which it recommends, are also in the true interest of each individual,” (EPM 9.2.16). Furthermore, Hume issues a promissory note: that *his own theory will succeed in establishing the connection between true interest and morality*: “The peculiar advantage of the foregoing system seems to be, that it furnishes proper mediums for that purpose.” Ending his inquiry by suggesting that his own theory fails to live up to that promise would be ending on a sour note, indeed. I think we can do better by Hume. In particular, I think there is some evidence to establish that the sensible knave has a prudential reason—an “interested obligation”—to behave in accordance with justice. Even if this fails, however, though the sensible knave may not have an *interested* obligation (i.e., an obligation that conforms to the demands of prudence) to justice, he has a normative obligation nonetheless (which is more than can be admitted on a normative internalist reading).

The key passage is Hume’s concluding paragraph of EPM 9.2. This passage is often neglected, to Hume’s disadvantage. (In particular, Jason Baldwin suggests that this paragraph reduces those who would argue with the sensible knave to “stammering about the beauty of a clear conscience.”²⁴ As we will see, this neglect is unfortunate and unwarranted.) Hume begins this paragraph by conceding that though the deck is stacked against free riders, it is possible that a dishonest or unjust character might succeed in hiding his dishonesty or injustice from public view. But, for Hume, this is not the end of the story. Hume writes: “But were they ever so secret and successful, the honest man, if he has any tincture of philosophy, or even common observation and reflection, will discover that they themselves are, in the end, the greatest dupes, and have sacrificed the invaluable enjoyment of a character, with themselves at least, for the acquisition of worthless toys and gewgaws,” (EPM 9.2.25). What is Hume trying to establish here? Given

²⁴Jason Baldwin, “Hume’s Knave and the Interests of Justice” in *Journal of the History of Philosophy* 42 (2004), 284. It is worth noting that Baldwin himself finds that this final paragraph suits Hume’s purposes as Baldwin sees them—that the sensible knave is merely an exception to the general rule that most people are just and honest. Unfortunately, this interpretation does not serve Hume’s purposes as Hume sees them at 9.2.16.

the structure of the rebuke Hume offers the sensible knave, Hume seeks to establish that the sensible knave has *bad taste*. Rather than desiring the invaluable enjoyment of a character, the sensible knave instead opts for the worthless gadgets that come along with sensible knavery. Hume is not stammering about the beauty of a clear conscience. He is noting that a clear conscience has beauty that the sensible knave cannot, or does not, see.

But recall that the standard of taste is normative. It provides reasons for action and desire. Thus there are perfectly good grounds for believing that Hume means to suggest that the sensible knave has a normative obligation to justice. In particular, Hume notices that one of the *explicit* requirements of the standard of taste is lacking in the sensible knave: comparison (OST 238). Hume continues: “How little is requisite to supply the *necessities* of nature? And in a view to *pleasure*, what comparison between the unbought satisfaction of conversation, society, study, even health and the common beauties of nature, but above all the peaceful reflection on one’s own conduct; what comparison, I say, between these and the feverish, empty amusements of luxury and expense?” (EPM 9.2.25). The rebuke of the sensible knave takes on precisely the form of the rebuke of the “bad critic” when it comes to the process of comparison: “A man, who has had no opportunity of comparing the different kinds of beauty, is indeed totally unqualified to pronounce an opinion with regard to any object presented to him,” (OST 238). But if this is the case, and the standard of taste provides reasons for action, we should agree that the sensible knave has a reason to behave as the “honest man” would. The honest man makes and understands the comparison between the “peaceful reflection on one’s own conduct” and the “empty amusements of luxury.” Indeed, the sensible knave has a reason to so conform, *though he has no motivation to do so*.

Thus I think it is clear that Hume maintains that the sensible knave has an obligation to justice. But does he have an *interested* obligation? Is justice in his *true interests*? The answer, I think, is yes. Many systems of norms are potentially reason-giving. Hume ties morality to the standard of taste, but there is no reason why he shouldn’t also tie *prudence* to the standard of taste. Indeed, this appears to be a natural reading of Hume’s rebuke of the sensible knave. He criticizes the sensible knave for preferring the wrong things, things that are not in his interest. Recall Hume’s introduction to the sensible knave passage. Hume writes: “Treating vice with the greatest candour, and making it all possible concessions, we must acknowledge, that there is not, in any instance, the smallest pretext for giving it preference above virtue, with a view to self-interest; except, perhaps, in the case of justice, where a man, taking things in a certain light, may often seem to

be a loser by his integrity,” (EPM 9.2.22). This passage might be read as follows: justice is an exception to the general rule that virtue is in the interest of the virtuous. But this reading is not correct. Hume’s qualifying terms “seems” and “taking things in a certain light” are crucial here. He goes on to suggest that the *reason* it “seems” as though the sensible knave might advance his interests by avoiding justice is because the sensible knave realizes that “an act of iniquity and infidelity will make a considerable addition to his fortune.” *But this is precisely what Hume criticizes in the ultimate paragraph of EPM 9.* The “certain light” in which the sensible knave must be taken is the “light” that Hume goes on to criticize, i.e., that adding to one’s fortune is truly better. Though it *seems* as though the sensible knave wins because he can make a considerable addition to his fortune, this addition is only chimerically in his interest, when things are taken in the wrong “light”.

EPM 9.2 is best read, it seems to me, by noting that though justice is not found in the prudential motivations of all individuals, it *is* found in the motivations of individuals who, like the “honest man”, are properly motivated by the joys of character. And if so, Hume’s promissory note at EPM 9.2.16 is vindicated. Hume’s account establishes an interested obligation to justice because *morality and prudence coincide*.²⁵ Prudence is not based simply on what would motivate a given agent. Prudence is given in just the same way morality is: by the verdicts of the standard of taste. Those who possess the standard of taste will understand that the “invaluable enjoyment of a character” is in their “true interests.” This understanding need not be moralized; the standard of taste may demand moral behavior, but only insofar as such behavior is required for the joys that accompany it—the enjoyment of a character, and the contemplation of one’s clear conscience. Hence, being necessary for these “higher” joys, justice is prudential. And here we have Hume’s ultimate rejection of normative internalism. Not only does he reject internalism when it comes to moral reasons—reasons that apply to us given our moral obligations. He also rejects internalism when it comes to our prudential reasons. For Hume, the standard of taste, not one’s

²⁵This reading of Hume is also advanced by Gerald Postema in “Hume’s Response to the Sensible Knave” in *History of Philosophy Quarterly* 5 (1988). Postema’s reading is to some degree different than my own, however: I claim that justice is in the interest of the knave because the various joys of character motivate those possessed of the proper sensibilities. Postema claims that, for Hume, justice is in the interest of the knave given the knave’s current motivations: as the knave seeks the esteem of others and himself. As Postema rightly notes, this move is not particularly plausible. The good news is that there is another interpretation that puts Hume’s opening declaration of EPM 9.2 on firmer footing: it is not the knave’s esteem of the joys of character, but rather the “honest man’s” esteem of the joys of character that entail that justice is in the knave’s true interests.

own motivations, governs one's prudential reasons. In the case of the sensible knave, he has reason to develop and live according to the higher joys of character, conscience, and sociability, for which justice is required.

Not so fast. My reading might be thought to fly in the face of Hume's declarations, in the *Enquiry* and the *Treatise*, that acts of justice can run counter to the agent's interests. There are several representative passages, but I'll confine myself to the following, from the *Treatise*:

Nor is every single act of justice, consider'd apart, more conducive to private interest, than to public; and 'tis easily conceiv'd how a man may impoverish himself by a signal instance of integrity, and have reason to wish, that with regard to that single act, the laws of justice were for a moment suspended in the universe. (T 3.2.2.22)

Here it appears clear that Hume intends justice, in particular occasions, to make people worse-off. Hence, justice cannot, in every instance, be in our "true interests". Hence, the sensible knave cannot be solved by suggesting that justice, rather than injustice, is in the interests of the knave.

The key is the contrast class. Saying that justice can sometimes make people worse-off does not entail that they would do better by injustice. Rather, Hume argues that justice can "impoverish" people—a just act at time t can make people worse-off than they were at $t-1$. As Hume writes at EPM App 3.7, it is impossible for justice to prevent "all particular hardships, or make beneficial consequences result from every individual case." But this just means that it is impossible that justice will make all people better-off at t than they were at $t-1$. All this entails is that not every act of justice will make you result in "beneficial consequences"—meaning consequences that make you better off than you were. Everything Hume says here is compatible with suggesting that *injustice* would make people much worse-off than acting in accordance with justice. Furthermore, Hume's claim that a person might have "reason to wish, that with regard to that single act, the laws of justice were for a moment suspended in the universe" is particularly telling. In considering the impoverished just man, Hume does not say that the just man wishes that he *would have acted unjustly*. Rather, Hume suggests that this person instead has reason to wish that the laws of justice were "suspended"—meaning, of course, that the seemingly beneficial action (perhaps one that might have increased his fortune) would have had at least neutral moral valence. But why should the just man wish that the laws of justice were suspended, rather than, simply, that he had acted differently? Plausibly, because—though it might increase one's fortune—the unjust

act is only chimerically beneficial *given* its injustice.²⁶ And this reading is supported by Hume’s declarations concerning the sensible knave. At EPM 9.2.23, Hume writes that all “ingenious natures” understand that character is “requisite” for happiness. But that justice is “requisite” for happiness does not entail that a just act is *sufficient* for happiness; nor does it imply that justice will always benefit the just. Rather, justice is in the interests of the virtuous, because it is *necessary* for those interests, not because (implausibly) it always makes one “better-off”. My claims are compatible with everything Hume writes at T 3.2.2.22, and EPM App 3.7.²⁷

Though there may be many reasons for rejecting this reading of the sensible knave, it has one manifest virtue: it allows Hume to accomplish his own goals, and allows him to fulfill a promissory note at EPM 9.2.16 that we have every reason to believe he intended to fulfill. However, in order to embrace these results, we must reject the claim that Hume is a normative internalist. No matter. This claim *should* be rejected.

5. Charity

Is this view charitable to Hume? As I suggested before, Hume’s attitude toward the moral domain is certainly plausible: Hume believes morality is normative, and morality is derived from the verdicts of the standard of taste. This conforms to our considered judgments about the normative force of morality, generally. But the following thoughts might motivate a negative answer in the prudential domain. For instance, imagine that Guinness is preferred to PBR by those with suitably delicate sensibilities. Imagine,

²⁶This reading is cemented by the sentence immediately following the quote given above: “But however single acts of justice may be contrary, either to public or private interest, ’tis certain, that the whole plan or scheme is highly conducive, or indeed absolutely requisite, both to the support of society, and the well-being of every individual.” Here Hume moves from the *diachronic* comparison (whether x is better-off at t than at $t-1$) to the *counterfactual* comparison (whether x is better-off having behaved unjustly rather than justly). He suggests that, no matter how injurious a particular act of justice might be to the interests of someone when it comes to the diachronic comparison (which establishes that justice can “impoverish” both man and society), justice always wins in the *counterfactual* comparison, because it is always “requisite” for the “support of society, and the well-being of every individual.” I read this as marching in lock-step with my reading of Hume’s declarations at EPM 9.2.23. But, at the very least, it is *not incompatible* with that reading.

²⁷EPM App 3.8 suggests that an agent might, out of “self-love” prefer an act of injustice. But self-love, for Hume, just is our selfish *motivations*. Merely because, for instance, sensible knaves will prefer injustice out of self-love does not entail that injustice is in their “true interests”.

however, that I simply despise Guinness. I don't possess the delicate sentiments. Surely, we should like to say, I don't have a reason to drink Guinness *now*. But if prudential reasons are tied to the standard of taste, this verdict appears blocked. Furthermore, though the sensible, delicate critic might prefer Antonioni's *L'Avventura*, I, in my benighted state, surely do not have reason to see it rather than McTiernan's *Die Hard*.²⁸

Hume can accept these views. The question in these cases is not so much *that* the standard of taste is normative, but rather *how* it is normative. Take the PBR fan. Though Hume is committed to the claim that the verdicts of the standard of taste are the source of prudential reasons, he needn't be committed to the claim that, whatever one's sensibilities are *now*, one is prudentially better-off by doing whatever it is that delicate critics do. But it also does not mean that Hume must say that one has a reason to follow one's motivations in such cases. At the most general level, prudential reasons bear a *promoting* relation to the good: I have a prudential reason to promote whatever is in my own best interests. In the case of the PBR fan, that means *developing* one's taste such that one can enjoy the finer pleasures, desired by the delicate critic, of Guinness rather than PBR.²⁹ This might yield a more specific reason, not to drink a Guinness immediately, but to skip the PBR and try to appreciate something more moderate; a Newcastle, say, working one's way up. (In a sense, the sensible knave is a far simpler case. For Hume, one has a prudential reason to promote one's true happiness—the "enjoyment of a character, with themselves at least". But in order to promote one's interest in this domain, one *must* act virtuously. As Hume says, justice is "required" for one's "true happiness" as identified by the "honest man": the joys of character, conscience, and sociability.) Similarly in the case of the fan of action flicks. In my benighted state, appreciating *L'Avventura* is simply unavailable. Nevertheless, despite one's motivations, one has a reason to avoid *Die Hard* and go for something that would develop one's sensibilities to a greater degree—say, *Vertigo*—working one's way to the more refined joys of a proper aesthetic sensibility. It is important to note, however, that in no case does Hume temper these reasons based on an agent's subjective motivational set. I regard this view as plausible. Or, at the very least, plausible *enough*.

²⁸Thanks to an anonymous reviewer for raising this worry.

²⁹In this way, Hume's account of the standard of taste might bear a striking similarity to Mill's discussion of competent judges in *Utilitarianism* II. I think this connection is not to be shied away from, and there are good reasons for believing that their views on value bear not only a second-order similarity, but a first-order similarity as well. This, however, is a topic for another time.

6. Conclusion

I have argued that Hume is not a normative internalist in the way he is often taken to be. Passages in the *Treatise* that seem to suggest normative internalism are inconclusive at best. Furthermore, an examination of the connections between “Of the Standard of Taste” and the *Enquiry* revealed that Hume holds that the standard outlined in his essay provides reasons even for those who fail to have the proper motivations. This reading also allows us to solve significant puzzles concerning Hume’s treatment of the sensible knave. This, it seems to me, is good reason to accept this reading. Hume is no nihilist about practical reasons. Though closer to the truth, he is no normative internalist, either.

There are lessons for contemporary ethics that we can and should draw from Hume’s rejection of normative internalism. The first, and least important in my view, is that contemporary internalists should think twice before referring to themselves as “Humeans”. (This lesson is least important because most have taken pains to avoid ascribing any particular view to Hume, though often because many have falsely assumed Hume a nihilist.) The second, and more important, is that Hume’s position illustrates one way of taking seriously the claim that “reason is ... the slave of the passions” without insisting that “reasons are based, ultimately, on subjective, contingent, conative states of the agent.”³⁰ “Reason”—even in the sense of a practical reason—is the slave of the passions for Hume; but it is not the slave of whatever our motivations happen to be. It is the slave of a *subset* of passions, passions that are possessed only by those who are properly attuned to the standard of passion itself.

Third, and I think most interestingly, Hume’s theory can interpret one plausible motivation for normative internalism in a novel and underexplored way. Leaving aside overall plausibility, at least one primary motivation for normative internalism is its built-in link between practical reasons and motivations. It would be hard for any theory to explain why reasons *typically* motivate agents unless there is at least a reasonably tight connection between practical rationality and antecedent motivation. Hubin, for instance, writes:

When [internal] 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

³⁰Hubin, “What’s Special about Humeanism”, *op. cit.*, 30.

by these 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. . . 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.³¹

Hume’s considered view allows that an agent, such as the sensible knave, will not “care” about the evaluative perspective that entails that he has reason to behave justly rather than unjustly. But in defending a connection between reason and motivation, Hume needn’t retreat to the near-empty suggestion defended by, e.g., Korsgaard that the only necessary connection between reason and motivation is the motivation of the *rational* person. Rather, Hume can insist that reasons do, in fact, correspond to antecedent motivations. But this link to motivation is mediated by a form of *perfectionism*. For Hume, reasons correspond not to the agent’s concern, but rather to *human concern*: what humans care about when they are free from the defects that distort their motivational sentiments away from those sentiments found in human nature *itself*. Hume believes that there is such a “sentimental nature”, and that this nature will be displayed in anyone with a lack of prejudice, sufficient delicacy, and sufficient training—maybe we agree, maybe we don’t. Hume also believes that those who do not react positively to the verdicts of the standard of taste, like the sensible knave, are few and far between³² (and hence that practical reasons will “typically” motivate)—maybe we agree, maybe we don’t. But for those who believe that normative internalism ties practical reasons too closely to an agent’s antecedent motivation (with the subsequent lack of power to rationally criticize desires from an external point of view), and for those who believe that rationalists leave too much distance between practical rationality and desire, Hume’s view might be a welcome *via media*. At the very least, Hume’s “sentimentalist perfectionism”, as it might be called, merits more investigation; investigation that has been neglected, I conjecture, given the mistaken association between Hume and normative internalism.

³¹Hubin, “What’s Special about Humeanism”, 39-40.

³²See, for instance, OST 242.