



Reply to Radcliffe and Garrett

Rachel Cohon

Hume Studies Volume 34, Number 2, (2008) pp. 277–288.

Your use of the HUME STUDIES archive indicates your acceptance of HUME STUDIES' Terms and Conditions of Use, available at

<http://www.humesociety.org/hs/about/terms.html>.

HUME STUDIES' 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 HUME STUDIES archive only for your personal, non-commercial use.

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

For more information on HUME STUDIES contact

humestudies-info@humesociety.org

<http://www.humesociety.org/hs/>

Reply to Radcliffe and Garrett

RACHEL COHON

I thank both my critics for their praise, their searching comments and objections, and their careful attention to my book. In the very short time allotted to respond to them both, I will address their objections in an integrated way, following the order of my book.

Both Elizabeth Radcliffe and Don Garrett protest that for the last twenty years the noncognitivist reading has not dominated Hume scholarship in the way that I suggest when I include it in the common reading of Hume's metaethics. In the book I admit that noncognitivism is not as popular among experts as the other two elements of the common reading, and I discuss the alternatives to it that have been proposed. But most of those who offer such alternatives read Hume as saying that beliefs alone cannot cause passion or action, and that moral judgments alone can and do. These two claims together entail that moral judgments are not beliefs, which is hard to distinguish from noncognitivism or non-propositionalism. So the noncognitivist reading is implicit in more interpretations than it seems.

I try hard in the book to look afresh at Hume's moral psychology, and to resist the pull of the common reading myself in interpreting it. In order to respond to Professor Radcliffe's and Professor Garrett's objections, I must explain a little more of what I do there.

Hume says that pleasure and pain are the sources of most of our motivating passions or affections. (By "motivating passions" I mean those passions that proximately cause action.) These passions are caused not only by the agent's occurrent pleasure and pain, but also by pain and pleasure "conceiv'd merely in idea, and . . . consider'd as to exist in any future period of time" (T 2.3.9.2; SBN 438).¹ In

Rachel Cohon is Associate Professor of Philosophy at The University at Albany, State University of New York, 1400 Washington Avenue, Albany, New York 12222, U.S.A. Email: rcohon@albany.edu.

particular, *beliefs* about one's own future pain or pleasure (I call these *hedonic beliefs*) yield passions that influence the will. (See also T 1.3.10.2; SBN 118–19.) According to the common reading, the only role of a belief in influencing the will is to provide information that directs some already-present passion toward its object. If this were true, then when Hume says that beliefs about pleasure and pain influence the will, he must mean that there is a desire for pleasure and an aversion to pain already in the mind to be directed by such beliefs. Garrett mentioned several ways in which I argue against this interpretation. A further point I make is this one: Hume says that when an idea of pain (for example) becomes enlivened to the point of being a belief, it comes to have the same influence on the passions and the will as does an *impression* of pain, though in a lesser degree (T 1.3.10.3; SBN 119–20). Now, an occurrent feeling of pain when I touch a hot kettle does not merely provide me with information about how to avoid the object of an independent aversion to pain; rather, it generates an aversion. Hume's hedonism, I think, appeals to the inherent aversive power of felt pain and attractive power of felt pleasure. So if the *belief* that touching a hot kettle *will* cause me pain has the same influence on my passions as a present impression of pain, then the belief, too, does not merely provide me with information about the object of an independent passion, but actually causes me to form a new passion.

I also argue that Hume understands reason as a process or mental activity, reasoning. He describes it as one of comparing perceptions and discerning relations between them, finding some ideas to correspond with other impressions (that is, to be true). As an empiricist, he cannot identify the faculty of reason first and then figure out what functions it performs; he must start with what we can observe, which is the comparing and linking of our ideas and finding evident their relations to one another. I propose that when Hume says something is produced or made by reason alone, he means it is produced by a process purely of this type, without the aid of another type of process. Since reasoning is simply comparing perceptions and finding true, all that can be produced by this method alone is a lively idea (or more precisely, the enlivening of an idea), not a new impression.

Some (though not all) hedonic beliefs are produced by causal reasoning alone—by comparing a series of believed ideas of past conjunctions with a present impression to find a further idea to be true. And hedonic beliefs, including those produced by reasoning alone, cause new motivating passions. But I argue that this does not entail that such passions are caused by reason alone. Hume does not say much about the process of passion-formation, but it is fairly clear that it is not one of reasoning. Passion-formation is the same sort of process that goes on when a felt pain causes an aversion; and that is not a process of comparing perceptions and finding true. If I am right that the expression “reason alone” means “by a reasoning process, without another kind of process,” then it follows that motivating passions are never produced by reason alone—they require another type of process as well.

Radcliffe discussed one worry about this. If, as I acknowledge in my interpretation, causal reasoning alone produces some hedonic beliefs, and these in turn cause motivating passions, it may seem to follow that causal reasoning alone produces motivating passions. This inference depends on the assumption that the relation “is the cause of” (or “is the sole cause of”) is transitive: if A is the cause (or sole cause) of B, and B is the cause of C, then A is the (sole) cause of C. I doubt that this principle of transitivity holds even for the causation of material objects or events; but I leave that issue aside. I argue that it definitely does not hold for processes. My examples of the manuscript and statue are meant as counterexamples, to show that this is not a valid rule of inference. Consequently, even given the causal power of hedonic beliefs, Hume can claim, without inconsistency, that passions are not caused by reasoning alone, since they require a further process after the reasoning is complete. Garrett is right that my objections to the transitivity principle are not Hume’s. However, they are not intended to show that Hume held any particular position about the transitivity of causation. My intent is to show that if Hume indeed understands reasoning and the production of passions as I say he does, in terms of processes, an objection based on the transitivity of causation would be misguided, because based on an assumption that is false in this context.²

Radcliffe argues that “when a belief is acquired by a piece of reasoning, that belief is an essential component of that line of reasoning or of the reasoning process” (ER, 269).³ So, she says, “it’s hard to see how any effect of a conclusion on the investigator is not also an effect of the reasoning from which it was derived” (ibid.). I gather she means here that if reasoning produces a hedonic belief which in turn produces a motivating passion, since the belief is *part* of the reasoning, the reasoning is the proximal and indeed sole cause of any passion caused by the belief, something Hume denies.

Of course the conclusion of a line of reasoning is an essential component of the *argument* that we construct when reasoning, if this is thought of as a series of propositions. After all, it’s the *conclusion*. But it sounds odd to me to say that the belief that results from the reasoning process is a component part of the *process*. Consider analogies with specific mental processes. If I come to form a belief by a process of adding numbers, my conclusion (say, that I have twenty-three dollars) is not part of the mental process of addition. Or take a different (Humean) mental process: acquiring a belief by sympathy with a charismatic person who holds it. My resulting belief is not part of the process of sympathetic transfer. There is a difference between a process and that to which it is applied. A process is repeatable with different material.

In any case, even if the resultant belief *were* thought of as part of the reasoning process, the causation of the passion by the belief is not itself part of any process of comparing ideas and finding true. If by “reason alone” Hume means reasoning without another type of process, then the passion is not produced by reason

alone. I do not deny that reasoning contributes to the outcome; I deny only that the passion is produced solely by that process.

Radcliffe is concerned about how to individuate mental processes. But in fact my interpretation says a good deal about the nature of a reasoning process. Reasoning, for Hume, is not just any sequence of events in the mind that ends with a belief. It is a process of a particular type: the comparing of perceptions and finding an idea to be a true copy of another perception. Such a process can indeed be part of a longer sequence of events that includes action, but that does not make the action the product of a rational process “by itself.” Only part of that sequence is composed of an instance of reasoning.

I argue next that when Hume talks about morals and moral distinctions he has in mind a mental process or activity as well, that of distinguishing good from evil; and he asks whether we perform this act by reasoning or by another type of process. Thus he asks “*Whether 'tis by means of our ideas or impressions we distinguish betwixt vice and virtue*” (T 3.1.1.3; SBN 456, underscore added), and he characterizes the moral rationalists as those who “concur in the opinion, that morality, like truth, is discern'd merely by ideas, and by their juxta-position and comparison” (T 3.1.1.4; SBN 456–57). His conclusion in Argument M, I say, is that the process of moral discrimination is not itself a reasoning process.

Garrett asserts that on my interpretation the impotence of reason is a consequence of the definition of reason and not of any argument about reason's causal power; whereas Garrett holds that Hume's actual argument for the impotence of matter-of-fact reason is causal. Garrett is right that part of Hume's evidence for the inertia of reasoning alone is empirical. Hume makes the following observation in T 2.3.3.3 (SBN 414) and in the *Dissertation of the Passions*⁴: when neither of two objects individually provokes any passion, the discovery by reasoning that they are linked as cause and effect also will not provoke any. This provides some support for the inertia of reason. But Hume treats the inertia-of-reason claim as a necessary thesis (reason of itself *cannot* influence passion), and I think he offers conceptual grounds for it as well, that emerge, in part, from the Representation Argument.

Garrett further objects that I treat the production of action as the defining characteristic of the process of moral discrimination. I don't think I do treat action-generation as defining, and I do not mean to say that Hume provides any definition of moral discrimination in the course of constructing his versions of Argument M. The definition of moral discrimination comes at the end of 3.1.1 and in 3.1.2 in Hume's articulation of the moral sensing view, and it is a definition Hume cannot deploy earlier against the rationalists without begging the question against them. But I do interpret Hume as specifying the capacity to produce passion as a necessary feature of the process of distinguishing good from evil. Hume thinks even his opponents must grant that there is a necessary link between moral

discrimination and feeling—indeed, between moral discrimination and volition (see T 3.1.1.22; SBN 465–66). This might merely be a link of the following kind: if a process of a certain type *never* of itself produced a feeling, that process would not be moral discrimination. Since reasoning processes of themselves never do produce a new impression, they cannot be processes of moral discrimination. That still leaves open the definition of moral discrimination.⁵

Radcliffe grants that Hume may well be denying the identity of the two processes in Argument M; but she doubts that this could be all Hume means to establish. His rationalist opponents hold that reason produces representations of good and evil, and these move the will (either of themselves or by giving rise to other psychological states). So in opposing the moral rationalists, Hume also means to argue that no rationally-produced representations can move the will. She thinks it would not satisfy Hume to show that beliefs do not produce passions *by a process of inference*, for, as Radcliffe understands him, Hume intends to deny that beliefs can do so by any process whatsoever, or any combination of processes. I do not see that he must say this to oppose the moral rationalists as thoroughly as he would like. He argues that reason alone cannot *discern* moral good and evil; moral good and evil are not known by reasoning alone (by the comparison of ideas). His longest discussion in T 3.1.1 is in support of this. If he establishes this, then the rationalist claim that reason (alone) produces our *representations* of moral good and evil simply is not true. Reason alone does not give us this knowledge. So Hume gets off the train at that stop, with no need to show that rationally-generated beliefs cannot move us to act by any process whatsoever.

Garrett doubts whether Hume is really a moral anti-realist (of the sort who thinks moral properties in themselves are reaction-dependent relational properties). He grants that according to Hume our *concepts* of good and evil as well as our judgments of them depend on the existence of the moral sentiments. But he thinks there is room to allow that for Hume the *property* of vice in itself, construed as the “categorical causal basis for the impression” of vice (DG, 258), is independent of the existence of our moral sensibility. For my part, I find Hume to be extremely reluctant to make any metaphysical commitments one way or the other; which is one reason why I back off, in part, from the claim that he is a moral anti-realist, and allow the possibility that the most we can say is that he is not a moral realist. (Probably Garrett would challenge this description of Hume too.) However, Hume’s rhetoric in the moral *Enquiry* and “The Sceptic” certainly seems to commit him to the view that were there no human beings with sensibilities like ours, there would be no moral properties, which is incompatible with the claim that the moral property itself is independent of our moral sensibility. Recall what Hume says in his extended analogy between “moral beauty” (virtue) and the beauty of the circle: “Till such a spectator [someone with an intelligent mind, susceptible to the finer sensations] appear, there is nothing but a figure of such particular dimensions

and proportions; from his sentiments alone arise its elegance and beauty” (EPM App. 1.14; SBN 292).⁶ If Hume has any position on what vice in itself is, if he even thinks such a question makes sense, then given such passages, it is far more likely that he does not regard vice itself as anything independent of human reactions, but rather as the power to cause sentimental responses in observers.

Turning to the common point of view, I argue that we experience a felt (if sometimes weak) sentiment whenever we imagine ourselves to occupy it. I also argue that we learn to adopt it while pursuing another purpose: to predict what practical effects a person’s traits will have on ourselves and others. From a distance someone’s irritable temper may seem amusing, or someone’s kindness a sign of weakness. But if I make the effort to imagine the impact of these traits on those directly affected by them, I recognize that her irritability is likely to cause hatred and shame, and his kindness is likely to cause love and pride. This information saves me and others much grief.

Radcliffe and Garrett are concerned that in so arguing I make room in Hume’s ethics for observer-independent moral qualities; for on my reading we figure out which companions will make our lives better based on their characters. Now, I assert all along that qualities of mind such as irritable temper and kindness *are* observer-independent qualities. It is their goodness or evil, their being virtuous or vicious, that depend on the existence of approval and disapproval. According to my account, we have an incentive to seek the common point of view (initially, at least) in order to learn whether a given trait in ourselves or another is likely to make our lives go well; but adopting this perspective, which enables us to imagine the pride, humility, love and hatred likely to be produced by a trait, also brings about an alteration of our situated feelings of approval and disapproval, making them genuinely moral feelings. Radcliffe thinks this implies that “it must be the *moral* qualities of others, and not just their factual qualities, that we need to get right in order to place our trust in the proper people,” and so people’s virtuousness and viciousness “must be real qualities independent of observers’ reactions to them” (ER, 274). But it does not follow from the process I describe that observers must get the moral qualities right in order to place their trust in the proper people. I do not need first to judge that irascibility is a *vice* in order to know that I would soon come to hate an irascible employer were I to become his subordinate. I need only to learn (or know enough in general to imagine) that his irascibility causes or would cause his current employees to hate him. To learn this, I must expand my attention beyond myself as I am now to include those affected by his trait. Thanks to sympathy, the same process that enables me to estimate these effects also transforms my situated approval or disapproval into a genuinely moral sentiment.

Garrett challenges my formulation of the puzzle about the virtue of honesty, and my proposed solution. He agrees that Hume is committed to my statement (1),

that honesty is a virtue. As for (2), that for every virtue there is in human nature some non-moral motive that is the object of approval, unlike me, Garrett thinks Hume consistently accepts a thesis only slightly weaker than that. Garrett denies that Hume holds (3), that there is no morally-approved motive of honest action distinct from the regard to honesty or morality (that is, distinct from the moral sentiment). By contrast, I think that Hume initially asserts (2) as something that “may be establish’d as an undoubted maxim,” but later relaxes it substantially. I will come back to that. I think Hume adheres to (3) all along: there is no morally-approved motive of honest action that is distinct from the moral sentiment. This is evident in Hume’s long enumeration of the possible motives for honest loan repayment which leads to the conclusion that the only motive that reliably suffices for this virtue is the sense of duty (T 3.2.1.9–16; SBN 479–83); and he makes the same claim explicitly about fidelity to promises, that “If we thought promises had no moral obligation, we never shou’d feel any inclination to observe them” (T 3.2.5.6; SBN 518). In the moral *Enquiry* it appears in his first reaction to the sensible knave (EPM 9.23; SBN 283). Neither discussion pertains to human beings in their pre-conventional state, so neither is restricted to non-artificial motives.⁷

As I read Hume, he treats self-interest as the motive to create the conventions of property at the start, and also as the initial motive to conform to them in a small community; however, the ultimate motive to conform—the one that is approved and accounted virtuous—is the “sense of honor or duty,” an augmented moral sentiment. (On reflection we approve this specialized feeling of moral approval or disapproval.) Garrett has a different interpretation: that Hume actually finds the non-moral motive to honesty that he seeks (though it is an artificial one). According to Garrett it is the policy of conforming to the rules of property as a system; it is distinct from self-interest (although adopted for self-interested reasons); and it is approved and therefore virtuous. Elsewhere Garrett describes it as the disposition to treat the rules of property as what Joseph Raz calls exclusionary reasons, reasons not to take other (countervailing) reasons for acting into account.⁸

Although I cannot elaborate on this here, I must say that I do not think the psychological state that Garrett describes qualifies as a *motive* of the sort that can constitute a Humean character trait. I also seriously doubt that *Hume* provides us with this appealing idea of a mental stance that persons can choose to adopt, and that, once adopted, is not subject to reassessment by specific considerations of self-interest. He certainly does not mention it as the non-moral motive that elicits our approval and so gives rise to our moral approval of just acts. The passage Garrett quotes about “fixing an inviolable law to [one]self” to conform to the rules of equity in fact describes what people do “after the opinion, *that a merit or demerit attends justice or injustice*, is once firmly establish’d among mankind” (T 3.2.2.27; SBN 501, underscore added); that is, after we already approve of justice, and so after the *moral* motive is in place.⁹

However, if I am right that Hume thinks we have no approved motive of honesty other than a moral one, I have painted him into a corner.¹⁰ I extricate him by saying he does not really accept (2) in his own voice but treats it as part of an inadequate commonsense conception of virtue. Garrett thinks my appeal to such a conception lacks textual support. (He calls it a desperate measure.) In response I would argue that there is strong textual evidence for one general point: Hume acknowledges a conception of virtue embedded in common sense which differs from his own philosophical conception. He talks about how our natural “partiality and unequal affection . . . not only have an influence on our . . . conduct, but even on our ideas of vice and virtue; so as to make us regard any remarkable transgression of such a degree of partiality . . . as vicious and immoral” (T 3.2.2.8; SBN 488). He refers to our “natural uncultivated ideas of morality” (ibid.). These remarks pertain to our state prior to the invention of property; but even afterwards, Hume says we find it difficult to reconcile the features of our artifice with our natural way of looking at things. For example, property ownership does not admit of degrees; but “in our common and negligent way of thinking . . . we find great difficulty to entertain that opinion, and do *secretly* embrace the contrary principle” (T 3.2.6.8: SBN 530). Even judges struggle with the legal requirement to assign a piece of property entirely to its rightful owner, because “half rights and obligations, which seem so natural in common life” have no place in their law courts (ibid.). Another aspect of the commonsense view appears with regard to fidelity to promises: we “feign” a special act of the mind in promising because we suppose fidelity to be a natural virtue (T 3.2.5.12; SBN 523). In common life, then, we have a conception of virtue and vice that suits the natural virtues but is in various ways at odds with the virtues of honesty and fidelity. That much is in the text.

I admit that Hume does not explicitly state something further, that one tenet of the commonsense conception is that every type of virtuous action must have an approved, non-moral motive. That is a piece of the picture that I have filled in. But I think it is continuous with the rest, and explains a great deal about honesty and the other artificial virtues. For example, it completes the explanation of why Hume insists that, as a consequence of our failure to realize that fidelity to promises is an artificial virtue, we suppose that what makes each promise morally binding is the performance of a special act of the mind. We pretend that there is such a mental act because we assume there must be some non-moral, approvable motive for every type of virtuous action, and given the fact that promising generates new moral obligation, the appropriate motive would have to be created with the making of the promise.

Once more, I thank Professors Radcliffe and Garrett for their penetrating and challenging comments, which for me have stimulated much thought. I am extremely fortunate to have such critics.

NOTES

1 References to “T” are to David Hume, *A Treatise of Human Nature*, vol. 1, ed. David F. Norton and Mary J. Norton (Oxford: Clarendon Press, 2007). References are to Book, part, section, and, where appropriate, paragraph. Where paragraph numbers are included, the accompanying references to “SBN” are to pages of Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch (Oxford: Clarendon Press, 1978).

2 I do something similar later. I claim that the content of a judgment that attributes moral goodness to a person is simply “that person has a property that feels like this,” and so a moral judgment is not a causal judgment; yet at the same time I allow that Hume may well think that moral qualities, metaphysically speaking, are powers to cause sentiments in observers. On the moral sensing view we do not predicate *causal* properties of people when we make moral judgments, since we typically judge that a person or trait is good simply by feeling it to be so, just as we typically judge an object to be blue by looking and seeing that it is. But the property “virtue” might be, in itself, a causal, relational property (a power), though we do not think of it that way. To see any inconsistency here, I argue, is to make the mistake of thinking we believe everything about the morning star that is in fact true of Venus. But I do not claim that Hume makes or is even aware of this argument, as though he had anticipated Kripke and Putnam; I merely say that if he does think moral goodness is in fact identical with a causal power, this does not in fact conflict with his view that our moral judgments are not causal in their content.

3 References to my critics’ papers are given as follows in the text and further endnotes below: references to Elizabeth S. Radcliffe, “Reason, Morality, and Hume’s ‘Active Principles’: Comments on Rachel Cohon’s *Hume’s Morality: Feeling and Fabrication*,” in the present volume, are indicated by “ER” in parentheses followed by a page number. References to Don Garrett, “Feeling and Fabrication: Rachel Cohon’s *Hume’s Morality*,” also in the present volume, are indicated by “DG” in parentheses followed by a page number.

4 *Dissertation of the Passions*, section 5, part 1. See David Hume, *Four Dissertations and Essays on Suicide and the Immortality of the Soul*, ed. John Immerwahr, John Valdimir Price, and James Fieser (South Bend, IN: Thoemmes Press, 1992 and 1995), 170.

5 Garrett offers a nice alternative to my construal of the conclusion of Argument M as denying the identity of processes. The faculty of reason by itself is insufficient to move the will, on Hume’s view, Garrett says; it needs passions as a separate faculty. However, there is no need to presuppose the presence of any passion *antecedent* to the belief (or independent of it); what is needed to move the will is only a significant contribution from passion once the belief is in place. This keeps the claim of the inertia of reason a causal one, and does not appeal to any controversial doctrines about the transitivity of causation. I have no objection to this proposal, as far as it goes. But it leaves us wondering why the faculty of reason should be insufficient by itself to move the will. Hume’s theory of causation and his empirical observation in the paragraph at T 2.3.3.3 (SBN 414) do not provide an answer. I think his conception of reason does provide one, and also some grounds for him to say (as he does) that reason alone *cannot* move the will.

6 References to “EPM” are to David Hume, *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 1998). References are to chapter and paragraph. The accompanying references to “SBN” indicate the corresponding page in David Hume, *Hume’s Inquiries*, ed. L. A. Selby-Bigge and P. H. Nidditch, 3rd ed. (Oxford: Oxford University Press, 1975).

7 Garrett, in his ingeniously-argued article “The First Motive to Justice: Hume’s Circle Argument Squared,” *Hume Studies* 33.2 (2007), (but, unfortunately for me, not actually published until September 2009), emphasizes the word “naturally” in the *Treatise* passage that reads as follows: “From all this it follows, that we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance” (T 3.2.1.17; SBN 483). “Naturally” did not appear in the original published *Treatise* but was added in Hume’s hand to his own copy of it, presumably for a future second edition that never came into being; consequently Nidditch added this word to the Selby-Bigge/Nidditch edition in 1978. Garrett claims this word shows that Hume meant to deny only that there is a natural motive to justice as contrasted with an artificial one; which leaves the possibility of an artificial but non-moral motive. I suspect that had Hume completed his revisions, he would not have added that word at that point, because its addition makes nonsense of the sentence, whether “natural” is taken to mean “non-artificial” or “non-moral.” The sentence has the form “We have no A but X,” which is equivalent to “We have exactly one A and it is X.” Thus, with the addition of “naturally” construed as a limitation on motives, the statement “we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance” says that we have exactly one natural motive for observing the laws of equity, namely, the very equity and merit of that observance (the moral motive); and so it entails that the moral motive to acts of equity is natural. This contradicts Hume’s expressed view, whether we construe “natural” to mean either non-artificial or non-moral.

Garrett suggests that the passage about fidelity to promises that I quote should similarly be read as though Hume had restricted his point to non-artificial motives (Garrett, “First Motive to Justice,” 278–81). That passage in full reads “Now, ’tis evident we have no motive leading us to the performance of promises, distinct from a sense of duty. If we thought promises had no moral obligation, we never shou’d feel any inclination to observe them” (T 3.2.5.6; SBN 518–19). Garret claims that the restriction to natural (that is, non-artificial) motives is implicit in context. But then once again it is difficult to make any sense of these sentences. Hume could sensibly say that we have no non-artificial motive of fulfilling promises, but not that we have none “distinct from the sense of duty,” for the sense of duty, on Hume’s account of this virtue, is not a non-artificial motive. Nor could he mean to say that if we thought promises had no moral obligation, we never should feel any non-artificial inclination to observe them; that is not a thesis he defends.

8 Garrett, “First Motive to Justice,” 276; Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1988), 124.

9 That the just person subscribes to a policy that displays such resistance to reassessment when facing current (dis)advantage is a plausible philosophical proposal independent of Hume scholarship, one that has been advocated by David Gauthier (“Reason and Maximization,” *Canadian Journal of Philosophy* 4 [1975]: 424–33; and *Morals by Agreement* [Oxford: Oxford University Press, 1986], chap. 6). (Gauthier thinks

that in becoming moral we choose to adopt a new pattern of practical reasoning. It is questionable whether anyone can adopt such a thing intentionally; but he seems right that moral individuals treat the norms they accept as reasons that outweigh or silence certain other sorts of reasons.) The general idea of a policy or plan that functions as a reason that resists reconsideration is developed in detail by Michael Bratman in *Intentions, Plans, and Practical Reasoning* (Cambridge, MA: Cambridge University Press, 1987). Stephen Darwall claims to see in Hume's *Treatise* the thesis that just persons treat the rules of justice "as authoritative, as in themselves giving reasons." *The British Moralists and the Internal 'Ought': 1640–1740* (Cambridge, UK: Cambridge University Press, 1995), 293. Darwall suggests that this disposition to regulate one's conduct by the rules of justice could qualify as the virtuous, non-moral motivational state that allows Hume to escape the circle. However, Darwall thinks Hume does not consistently identify this as the motive of just individuals, since in the *Treatise* Hume often seems to claim that just action always serves enlightened self-interest and could be motivated by that alone. Darwall further thinks the recognition of a motive consisting of a disposition to be governed by norms is incompatible with Hume's official hedonistic moral psychology; and he also notes Hume's emphasis on the moral motive as the characteristic motive of just action. (See Darwall, *The British Moralists*, chap. 10.)

I think the idea that becoming a just person consists of "making oneself into the kind of person who can consistently follow honesty as a policy" (DG, 12) is one that is rooted in much philosophical thinking about practical reason since Hume, including the Kantian idea of self-legislation, and I do not see this kind of deliberative and deliberately-chosen structure in Hume's understanding of character. If Hume thought it possible to make certain rules into considerations that would move us more powerfully than would particular opportunities to promote our own interests, it is difficult to explain why he thinks we are in fact extremely likely to fail to conform to advantageous systems of rules even in the face of lesser temptations. Hume thinks that in a large society we will often be tempted to act unjustly, in particular because of our susceptibility to the allure of immediate over long-term advantage; and he describes this particular disposition as a feature of our nature that "study and reflection . . . the advice of friends . . . and repeated resolution" cannot alter (T 3.2.7.5; SBN 536–37). If we had the capacity to make ourselves into the kind of person who regards a rule as a reason that trumps personal advantage, we could similarly make ourselves into the kind of person who regards a rule as a reason that trumps near-term gain—and who could act on that reason. But apparently we cannot. Fortunately there are other ways to engineer motives, whether to conform to justice or to an exercise program; we can manipulate conditions to provide ourselves with immediate incentives. Hume exploits this option. As I argue in chapter 8 of *Hume's Morality*, an advantage of the sense of duty as a motive to just action (and also to acts of allegiance to government) is that it is immune to the siren song of proximal benefit. For Hume, what excites us to conform to the rules of justice or the laws of our rulers is not a commitment to a policy that we do not reconsider at times of temptation, but the expectation of immediate unpleasantness, whether in the form of civil punishment or of acute self-disapproval, if we do not.

10 Garrett also claims that Hume does not even assert my claim (2), that for every virtue there is in human nature some non-moral motive that is the object of approval, and that I mistakenly think he does so because I conflate two different claims that he does make. These are: (a) The first virtuous motive that insures the merit of any action cannot be a regard to the virtue of that action, and (b) No action can be virtuous

unless there is in human nature some motive to produce it distinct from the sense of its morality. The difference that Garrett mentions between these is that Hume does not say that the same motive is both the original non-moral motive *in human nature* and the first virtuous non-moral motive. I am not sure I understand exactly what difference Garrett is finding here, but my guess is that he thinks (a) says nothing about whether the motive is “in human nature” whereas (b) specifies a motive in human nature. So while Hume does say that the first virtuous motive cannot be a motive of moral approval or disapproval, he does not say that the first virtuous motive must be a natural one. This leaves open the possibility of an artificial but nonmoral motive of the kind Garrett takes Hume to propose. Now since Hume derives (b) from (a), as Garrett acknowledges, I have to suppose that Hume intends a restriction to human nature to be implicit in (a) in the first place. So Hume does commit himself (or at least does seem to commit himself) to (2) as I state it. But I guess here Garrett is really agreeing with me that Hume relaxes (2). It’s just that Garrett thinks the thesis Hume calls “an undoubted maxim” is weaker than (2), and he thinks Hume never gives *that* up, since Garrett claims to find in Hume’s account an artificial yet non-moral motive for acts of justice. I think that the particular motive Garrett claims to find is not described in the text; Hume does not identify a non-moral motive of honest actions, artificial or non-artificial, that elicits approval. So I do not think the specification in (2) that the motive be “in human nature” matters much. I could have left it out. The only motive I see in Hume’s account that is a general or universal motive to honest action that is approved is a moral one, the sense of duty.