



Hume on Motivating Sentiments, the General Point of View, and the Inculcation of "Morality"

Elizabeth S. Radcliffe

Hume Studies Volume XX, Number 1 (April, 1994) 37-58.

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/>

Hume on Motivating Sentiments, the General Point of View, and the Inculcation of “Morality”

ELIZABETH S. RADCLIFFE

I. A Puzzle

Two very different interpretations of Hume’s moral theory are quite common in the literature: the “sentimentalist” interpretation, as I will call it, and the “ideal observer” reading. The sentimentalist interpretation has it that moral judgments are based on, or possibly even identical to, actual human feelings. Further, on this reading of Hume, virtue and vice themselves depend respectively on what produces pleasure or pain *in us* as normal human beings when we reflect on certain motives to action. The ideal observer reading, on the other hand, has it that, for Hume, moral distinctions are based on the hypothetical feelings of an ideal spectator—one who is, among other things, fully informed, entirely objective and not self-interested. Thus, virtue and vice themselves do not depend on actual human sentiments, but on the projected sentiments of an ideal spectator.

Among those philosophers who have adopted the sentimentalist reading of Hume are Philippa Foot, Stephen Darwall (at one time), J. L. Mackie, and Simon Blackburn. Since Hume emphasizes that moral distinctions arise from “inside” us in some respect and that this internal ground gives them a motivational efficacy, the sentimentalist interpretation often emerges in these philosophers’ discussions of Hume on moral motivation. For instance, in his earlier view, Darwall writes:

Note here that the moral judgment, for Hume, is internal to the moral judge. Since he believed that there would be general agreement in moral sentiment within a society, a sentiment within the moral judge would, he thought, generally be present in the agent also, were he to be placed in the appropriate reflective circumstances.¹

And Foot, in commenting on Hume's solution to the "is/ought" problem, writes:

Hume thought he himself had hit on the perfect solution to the problem. The new element in a proposition about virtue was the reference to a special sentiment of approbation: nothing new in the object, but something in ourselves. At a blow, he seemed to have put an end to the hunt for mysterious extra properties, and also to have shown the necessary connexion between morality and the will. For the moral sentiment, the special feeling which we call approbation, was a pleasurable sentiment, by which we were inclined towards those actions whose contemplation gave rise to it.²

For Mackie and Blackburn, Hume is an "emotivist" of a sophisticated sort and holds that moral judgments express a sentiment (a version of sentimentalism) but also accommodate "objectivist" ways of speaking about morality.³

Gilbert Harman, John Rawls, and David A. J. Richards are among the philosophers who have classified Hume's theory as an "ideal observer" theory.⁴ For example, Harman writes:

...according to Hutcheson and Hume, for an action to be wrong is for the action to be such that it would displease normal observers under conditions ideal for reacting to actions....

The relevant experience does not have to be actual. It is the experience an observer *would* have under ideal conditions.⁵

And Rawls writes:

Consider the following definition reminiscent of Hume and Adam Smith. Something is right, a social system say, when an ideally rational and impartial spectator would approve of it from a general point of view should he possess all the relevant knowledge of the circumstances. A rightly ordered society is one meeting the approval of such an ideal spectator.⁶

The ideal observer reading seems to be supported by Hume's text as well, since Hume points out a concern common to sentiment-based moralities: that the intensity of one's feelings toward persons' actions and characters varies

with the psychological distance from those about whom one is thinking, even though the moral judgments themselves do not vary. Hume explains this phenomenon by saying that in order to avoid the "continual contradictions" that would be occasioned by differences in feelings among individuals regarding the same character,⁷ we make moral distinctions by first "fixing on some steady and general points of view" (T 581-82).⁸ To a casual reader, it appears that Hume's point here is that we *alter* our feelings so that when we deem a character virtuous or vicious, we do it on the basis of new feelings we experience when we put ourselves in an impartial frame of mind, rather than on the basis of our original sentiments. Hume asserts, however, that getting into the requisite frame of mind is at least problematic; so, if our moral distinctions are really based on feelings experienced from a disinterested perspective, then our moral judgments must typically be produced by reflection on the feelings we *would* have were we impartial in the given situation.

That Hume's theory can be interpreted in two such widely divergent ways—as a version of sentimentalism and as a version of the ideal observer approach—is symptomatic of a puzzle ensconced in Hume's theory. How can the ground of morality be internal and motivating when an inference to the feelings of a hypothetical spectator is typically necessary to get to genuine moral distinctions? Hypothetical feelings are certainly not internal to human beings, since they are not feelings at all: no one has them. In this paper, I will ultimately present an interpretation of Hume that reconciles the *prima facie* conflicting strains. In Part II, I will show how the puzzle specifically arises from a conjunction of views put forth in Hume's *Treatise*, all of which must be taken seriously in a Humean ethics. These views lead, respectively, to the sentimentalist interpretation and to the ideal observer interpretation. It is important to show that the puzzle does arise in this way, since this question has at least gone unacknowledged by many commentators, including those mentioned above. Before I begin Part II, then, I ask readers to hold their respective interpretations of Hume's text in abeyance as they consider the source of the puzzle I have described; these readings are, I suspect, *responses* to the problem I am considering and depend on disputed (although not necessarily unjustified) renderings of Hume's views in various passages. In Part III, I consider one possible proposal for explaining how Hume's theory consistently incorporates the opposing strains, namely, the suggestion that in moral education the inculcation of "morality" internalizes the sentiments of the ideal observer. There I show why I find this suggestion problematic. Finally, in Part IV I defend my own reading, one in which I attribute to Hume a sentimentalism that also accommodates the evidence supporting the ideal observer interpretation.

II. Textual Evidence

The Basis for the Sentimentalist Interpretation

The sentimentalist interpretation is most strongly supported by Hume's arguments involving his claims about the motivating force of morality. When Hume argues that reason by itself cannot be the origin of our moral "rules," he claims: (1) Morality motivates; (2) reason alone does not; (3) therefore, reason alone cannot be the source of our moral distinctions (T 457). For this argument to be valid, premise (1) must be read as the claim that morality *alone* motivates, for if it motivated in conjunction with a desire (sentiment) the individual already possesses, it would have no feature that makes it impossible to be derived from reason, and the argument would be invalid. This is so, since reason obviously can motivate in conjunction with a desire; on Hume's view, reason is instrumental in ascertaining the means to the objects of our desires. Thus, if morality only motivates when conjoined with a previously-existing desire, and reason does so as well, morality *could* very well be derived from reason. If reason alone cannot motivate, then it must be that some kind of impression or affection is required to influence action; moreover, if morality *does* motivate, then the necessary affection must be present in those cases when it does.

What is this necessary affection that must be present in the agent in order for morality to motivate? On the reading required by the above interpretation of the argument against the rationalists and apparently accepted by those who adopt the sentimentalist reading of Hume, it is the affection from which our moral distinctions are also derived. In other words, the explanation of the motivating force of morality, with "morality" understood as referring to the moral distinctions we make, lies in these distinctions' being derived from our own feelings of approval and disapproval (forms of pleasure and pain), which also motivate us. Since we can only be motivated by sentiments *we* experience, if moral sentiments motivate, they must be *our* sentiments (rather than, say, the sentiments of an remote observer).⁹

Now, the question arises whether this interpretation of Hume implies that competing theorists, most notably the rationalists, hold that the ground of morality is *external* to us. After all, the rationalist Samuel Clarke maintains that the thought that an action is right or fitting may by itself motivate, and the thought that an action is right is internal to the agent.¹⁰ The crucial difference, however, between Hume's view and a view like Clarke's lies in their accounts of what constitutes morality itself. Of course, the drawing of or acknowledgement of moral distinctions is internal to us on any view, whether it is purported that we make these distinctions by reason or by sentiment, but on a theory like Clarke's, it is not a feature internal to our constitution that constitutes morality—that is, determines its content. Rather, we *use* reason on his view to discover the relations of fitness given in the nature of things. On

Hume's view, we not only rely on our feelings to draw distinctions between virtue and vice, but whatever it is that our feelings discern when they discern morality depends on our constitution. This point is quite evident in the *Treatise* when, after concluding that morality cannot consist in any relations of ideas discoverable by reason, Hume completes his case against moral rationalism by maintaining that it cannot consist in any matter of fact "outside" of us. (Such matters of fact are also objects of reason, since we must infer their existence from our present perceptions.) He writes:

The vice [of a deliberate murder] entirely escapes you, as long as you consider the object [action]. You never can find it, till you turn your reflexion into your own breast, and find a sentiment of disapprobation, which arises in you, towards this action. Here is a fact; but 'tis the object of feeling, not of reason. It lies in yourself, not in the object.... Nothing can be more real, or concern us more, than our own sentiments of pleasure and uneasiness; and if these be favorable to virtue, and unfavorable to vice, no more can be requisite to the regulation of our conduct and behaviour. (T 468-69)

Even though this excerpt is part of a longer passage whose meaning is much disputed, any plausible interpretation has to admit that Hume's argument is designed to point out a distinctive feature of morality as opposed to facts in the external world: Morality is discovered only by turning our attention inward on ourselves. This is not the case for the relations with which Clarke is concerned; the faculty of discerning those relations—reason—is internal, but the attention of reason is focused elsewhere. Hume's two claims—that you do not find the vice when you consider the object you call "vicious" and that you *do* find the vice only when you turn your reflexion into yourself and there find a sentiment of disapprobation—make a point about the conditions necessary to the "discovery" of vice, conditions that do not obtain when we attend only to the properties of external objects, among which Hume includes the murderer's passions, motives, volitions, and thoughts, and conditions that do not obtain when we attend to relations between ideas.

Consequently, the evidence for a *sentimentalist* reading of Hume's moral theory derives from Hume's argument against the rationalists based on his views about the motivating effect of morality, along with some other prominent (even though somewhat enigmatic) observations about the "internality" of morality. It is one in which *actual* feelings of human beings are in some way the ultimate criteria of their moral distinctions and one in which, as a consequence of the internality of the criteria, the distinctions have a practical push.¹¹

The Basis for the Ideal Observer Reading

When Hume turns to his detailed description of moral judgments as derived from a moral sense (in *Treatise* III 3 i), he offers a reply to an objection directed at his sympathy-based account of moral judgments. Hume traces the moral sentiments to sympathy, and his analysis of sympathy relies on the associative principles of resemblance, contiguity, and causality. Sympathy, for Hume, is the conversion of the thought (idea) of another's passion into the experience of a passion itself. Our ability to do this varies with the differences in our relations to others: our spatial and temporal contiguity to other persons, the degree of resemblance others have to us, and whether there are relations of causality between ourselves and others (our being related as family members) (T 319-20). Other things being equal, I am likely to be more pleased by the good character of my old friend than by that of someone I barely know, even though you may have the opposite reaction if you know the latter person, but not the former; nevertheless, we may judge the two to be approximately the same in moral virtue. Since our moral distinctions aren't influenced by these relations, the question Hume must answer is, "How can our moral concepts be derived from our sympathetic feelings?"

Hume maintains in response that we assume some standard perspective, a general point of view, as that from which our moral distinctions issue, and that we do so, as I have noted already, because otherwise we would have difficulty conversing about actions, and because different individuals attending to their feelings from their subjective vantage points would run into conflicts with each other (T 581-82). For the purposes of this discussion, I will often speak of "the general point of view"—that is, in the singular—not because anything I will say depends on its being singular, but because after Hume introduces the notion, he writes in a way that gives an appearance of singularity, with no further explanation.¹²

So, what are the regulatory features of the general point of view (GPV), and how does it differ from the particular perspectives? First, we survey a person's character from the perspective of the person's "narrow circle"—that is, we sympathize with the agent, her family, her acquaintances, her associates, and so on (T 591). This means we approach "equally near" someone emotionally distant from us, such as a figure in history, as we do to someone emotionally close to us, such as a person present to us in space and time (T 582).¹³ Second, we do not sympathize with the actual effects of a person's character traits, but rather with the tendencies of those traits. This is so because circumstances can prevent the operation of one's character: "Virtue in rags is still virtue; and the love, which it procures, attends a man into a dungeon or desert, where virtue can no longer be exerted in action, and is lost to all the world" (T 584). To be sure, there are many difficulties in interpreting these points, but my concern here is to understand how we achieve the

perspective in which all of the things described take place, that is, *how* we make moral judgments.

What constitutes attaining the GPV is unclear, since Hume's reference to the way in which we adjust our moral judgments is ambiguous. This ambiguity is reflected in his describing the GPV in two ways: as a "method of correcting our sentiments, or at least of correcting our language, where the sentiments are more stubborn and inalterable" (T 582). The difference between these descriptions is significant for Hume's account of motivation.

On the first reading of the method, which I call "correction of the sentiments," we alter the conditions of perception so that our feelings are in fact different from what they would have been under our initial circumstances. In other words, we perform a psychological maneuver (or a series of such maneuvers) to put ourselves in a "disinterested" frame of mind.¹⁴ This process is analogous to a person's correcting the conditions for perception in a dimly-lit room, by putting a swatch of fabric under bright white light to determine whether it is navy blue or black. Just as the color perception is considered definitive of the color of the fabric, so, too, the feeling of approbation or disapprobation experienced in the proper frame of mind is indicative of morality.

On the second reading of the method, which I call "correction of the language," we alter our moral pronouncements without any change in the feelings themselves. On this description, we don't "correct" the psychological conditions to affect the sentiments we experience because we are ordinarily incapable of suppressing our initial inclinations. This is especially so when the agent whose character we are contemplating is very close to us in virtue of the three Humean relations. Rather, we understand that our sentiments are influenced by our particular perspectives, and we compensate for our relations to others by considering how we *would* feel when the influence of relations is eliminated. This process is analogous to the correction a colorblind person might undertake (if it is indeed possible for her to infer how an object would appear to a normally sighted person), since she is unable to change her color perceptions by altering internal or external conditions.

The preponderance of Hume's explicit references to the GPV in the *Treatise* are to correction of the language, rather than correction of the sentiments. Consider the following:

1) Hume writes,

In order, therefore, to prevent those continual *contradictions* and arrive at a more *stable* judgment of things, we fix on some *steady* and *general* points of view; and always, in our thoughts, place ourselves in them, whatever may be our present situation. (T 581-82)

Hume describes us as "in our thoughts" placing ourselves in the GPV, because our present emotional and psychological constitution apparently does not conform to the demands of generality.

2) Of Marcus Brutus, Hume says,

We know, that were we to approach equally near to that renown'd patriot [as we are to 'our servant'], he wou'd command a much higher degree of affection and admiration [than he actually elicits from us]. (T 582)

3) He repeats this very point two pages later when he writes,

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

4) Furthermore, Hume admits that these principles of correction to which he is referring

are not altogether efficacious, nor do our passions often correspond entirely to the present theory. 'Tis seldom men heartily love what lies at a distance from them, and what no way redounds to their particular benefit; as 'tis no less rare to meet with persons, who can pardon another any opposition he makes to their interest, however justifiable that opposition may be by the general rules of morality. (T 583)

5) A bit later, he states,

The passions do not always follow our corrections; but these corrections are sufficient to regulate our abstract notions, and are alone regarded, when we pronounce in general concerning the degrees of virtue and vice. (T 585)

6) Observations about the limitations of our ability to alter the sentiments are evident in *An Enquiry Concerning the Principles of Morals* as well. There Hume qualifies his remarks about the GPV, saying, for example,

The judgment corrects or endeavours to correct the appearance: But it is not able entirely to prevail over sentiment. (E 228, n. 1)¹⁵

7) He continues in the *Enquiry*,

...though the heart takes not part entirely with those general notions, nor regulates all its love and hatred, by the universal abstract differences of vice and virtue, without regard to self, or the persons with whom we are more intimately connected; yet, have these moral differences a considerable influence, and being sufficient at least for discourse, serve all our purposes in company, in the pulpit, on the theatre, and in the schools. (E 229)

Consequently, in those social situations which depend on discoursing about morality, we can use the public language of morality even though our own affections are not always changed by the virtue and vice we acknowledge.¹⁶

(8) The last quotation is also followed by a footnote that emphasizes the point that it is in our nature that “private connexions should prevail over universal views and considerations.” Hume continues,

Thus a small benefit done to ourselves, or our near friends, excites more lively sentiments of love and approbation than a great benefit done to a distant commonwealth: But still we know here, as in all the senses, to correct these inequalities by reflection, and retain a general standard of vice and virtue, founded chiefly on general usefulness. (E 229, n. 1)

Consequently, our sentiments do not correspond precisely to the moral standards we adopt.

Such passages have been used to support the ideal observer interpretation. Hume’s various comments seem to imply that moral observers reflect about how they would feel if they were situated, contrary to fact, in a standard perspective disregarding their particular circumstances and then correct their judgments accordingly. The ideal observer reading does not emerge without some argument, however, and while it is not necessary here to detail that argument, it is important to ask just what these passages should be taken to represent. More precisely, are these remarks representative for Hume of the typical case of moral judgment formation, or does he mention these *because* they’re *exceptions* to the typical case? If these are meant as exceptional cases, then the typical instance may very well be one where an observer experiences the moral sentiments that produce moral judgments—and the ideal observer interpretation is groundless.

From the very introduction of the GPV in *Treatise* III 3 i (“Of the origin of the natural virtues and vices”), Hume treats the corrections we make from the GPV as correction of the language. The passages above where he suggests that the sentiments themselves do not conform to the GPV are not incidental to that section, but form part of his explanation of what the GPV involves. Only

by restricting one's attention to III 1 ii ("Moral distinctions deriv'd from a moral sense") might one conclude that correction of the language is the exception rather than the norm, since it is that section in which Hume introduces the notion of a moral sense where our feelings are the source of our moral distinctions; given the point of that section, no correction of any sort is mentioned there. His point is to argue that since morality is not founded on reason alone, it must be founded in sentiment, and the centerpiece of that discussion is the argument from motivation, requiring that our moral judgments come directly from our sentiments; so, the development of the apparatus of moral judgment is left for later, in III.3.i, where the GPV is introduced.

Furthermore, the excerpts themselves offer evidence about the strength of the thesis that our sentiments don't change in accord with the GPV from which our judgments are made. The first three passages, two of which concern how we regard historical figures, along with the eighth, which concerns how we regard distant peoples, simply acknowledge that when we make moral judgments, "we reflect on how we *would* feel if we sympathized with the "narrow circle(s)" of the agent(s). In these excerpts Hume treats as routine our correcting the judgments (the language) when our feelings are still influenced by our particular relations to others. The fourth passage says explicitly that our passions do not often correspond to these principles of correction and that people "seldom" love what is distant from them; whereas the claims in the fifth, sixth, and seventh passages are a bit weaker, indicating that the passions sometimes do not follow the corrections or do not follow them entirely.

On the basis of these passages and where they occur, we are justified in attributing to Hume the view that adjusting our moral judgments according to the GPV is not the exception, but the norm. The adjustment may be greater in some cases than in others, depending on the relations in which we stand to the agent whose character we are considering. Sometimes our proximity to the agent may make it possible for us to sympathize with the narrow circle and base our judgments on our actual feelings. But it is not an exaggeration of Hume's view to think that the latter cases are not very common. Hume has the apparatus for explaining how we edge ourselves into a "more" GPV, but it is obvious from a detailed examination of our psychology that he is right to consider such occasions when we needn't adjust our judgments rare, as the following discussion illustrates.

For Hume, sympathy is defined as "the conversion of an idea into an impression by the force of imagination" (T 427; see also 317). So, we might broaden the scope of sympathy, on Hume's account, by bringing our imagination into play and increasing the vivacity of the ideas of the pleasures and pains of those who are unrelated to me by contiguity, resemblance, or causality. When we read about the anguish of a Mideastern refugee brought on by the actions of a dictator terrorizing select populations of a country, we

might muster the vivacity of an impression of displeasure as we imagine the hardships this person endures. If we can do this with respect to a large number of people affected by the dictator's action, then the overall emotional effect on us as spectators who have managed to move closer to a GPV is roughly indicative of the morality of the dictator's character. Furthermore, although Hume doesn't make it explicit, one effective way to extend our sympathies is to broaden our experience. For example, I may be able to sympathize with members of a Hindu culture more easily—something I find difficult to do at present—if I live among them for a while, share their way of life and engage in practices typical of people with their beliefs, or perhaps if I read more about them.

Nonetheless, even though I have the capability of broadening my sympathies in the ways Hume describes, I will not feel as intense displeasure toward a thirteenth century Genghis Khan as I might feel toward a twentieth century Hitler or even more likely feel toward a 1990s Saddam Hussein and certainly feel toward a rapist at large in my neighborhood. We are typically unable to intensify the force of our ideas by the imagination or broaden our sympathies by experience far enough to bring our actual feelings in line with the moral judgments we make.

Consequently, an interpretative question emerges from a conjunction of points put forth in separate places in Book III of the *Treatise*. Hume explains the motivational force of moral beliefs, implicit in Hume's simple argument against the rationalists, by locating the ground of morality inside of us—in our own feelings of approval and disapproval. This point leads to the sentimentalist interpretation. So, if Hume is serious that the feelings I find inside myself when I grasp virtue and vice are motivating, then it seems that the inference to the feelings of the GPV do not play a part in producing that motivating effect. That inference, however, is typically necessary to get to moral judgments. Hume's claims about the GPV indicate that moral judgments require in most cases a process of reflection by which we correct our language when our sentiments will not conform to the demands of impartiality.¹⁷ This point leads to the ideal observer interpretation. Thus, our actual motivating sentiments of approbation and disapprobation do not track our considered moral judgments precisely.¹⁸

Something needs to be explained. The interpretive challenge in Hume is to reconcile the claim that (a) our moral distinctions are directly based on feelings not identical to the feelings we in fact experience, with the claim that (b) the motivating force of morality comes from the internal nature of its grounds.¹⁹ In response, one can

- (1) deny that Hume meant (a);
- (2) interpret (a) in a way that makes it consistent with (b);
- (3) deny that Hume meant (b); or
- (4) interpret (b) in a way that makes it consistent with (a).

Option (1) can be easily dismissed on the grounds of the evidence I have given from Hume's text. This implies that an emotivist interpretation, identifying moral judgments with occurrent feelings, is insupportable. Option (2), to re-interpret (a) in some reconciling way, can also be set aside, I think; identity of feelings does not admit of degree, and the collection of passages from Hume above constitutes substantial evidence that the feelings we experience are not identical to those that would correspond directly to our moral judgments. I also find no way to take option (3) and deny that Hume meant (b), since motivation on his view always originates in passions, which are internal states of an agent. Just one strategy remains.²⁰

III. The Inculcation of Morality?

As I have argued earlier, on Hume's view the moral sense has two features: epistemic and motivational. That is, the impressions from which our ideas of morality are derived are purportedly also the feelings from which motives to action arise. However, that the two are not really identical is the source of the interpretative puzzle under discussion here. It is interesting that the fact we must correct the language of our moral judgments when our feelings don't conform to a general perspective—the moral point of view—is not problematic for making sense of the epistemological role. That is, Hume's theory is quite plausibly read as a theory about how the *genesis* of our basic moral concepts lies in our human nature without committing him to the claim that there is a one-to-one correspondence between each individual's judgments and that individual's *own* feelings. One of the points of Hume's references to general rules (scattered throughout all three books of the *Treatise*) upon which people rely when they form judgments is to explain how an individual can apply concepts derived from impressions she has never experienced herself.²¹ This is how we are able to communicate about morality "in company, in the pulpit, on the theatre, and in the schools," despite the fact that most of us do not routinely get ourselves into a state of mind that disregards our personal connections.

Consequently, that we must "project" ourselves into the GPV is one of the rules which we apply when we want to arrive at correct moral beliefs, but that people have to apply it does not mean that moral distinctions cannot ultimately be grounded in human feelings. As Hume himself argues, following such rules does not imply that the "approbation of moral qualities...is...deriv'd from reason, or any comparison of ideas" (T 583); "[s]uch corrections are common with regard to all the senses" (T 582). Just as I can describe to someone who has never heard Mozart how his music sounds and enable that person to have a notion of Mozart's style and converse with others about his music, likewise, individuals who by Humean standards have never felt genuine moral approbation can acquire moral notions through education,

training, and discussion, and can communicate perfectly well about them with others. In neither case does this possibility mean that the relevant concepts did not *originate* in direct perceptual experience. So, despite the fact that we rely on general rules to correct our moral pronouncements, it is justified to say that morality is in a *general* way founded on human feeling.

Can the same considerations allow Hume to account for the *motivating* character of morality?²² Hume is making an observation about human nature—that morality *typically* provides motives for persons, not that it does in some very unusual instances.²³ Any plausible interpretation of Hume, if it is to explain how people experience what counts as *moral* motivation, must explain this: How do people who generally do *not* experience the exact feelings upon which their moral assessments are based frequently find that those assessments furnish motives, when motives are *feelings*?

One possible reply is that these moral assessments, or beliefs, when acquired by education or conditioning rather than by our own sentiments, take on the force and vivacity of impressions and, as such, have the effect of impressions of reflection, which are passions (feelings). Hume explains the impact of education on the intensity of our beliefs when he compares this impact to that of the constant conjunction of causes and effects:

All those opinions and notions of things, to which we have been accustomed from our infancy, take such deep root that 'tis impossible for us, by all the powers of reason and experience to eradicate them; and this habit not only approaches in its influence, but even on many occasions prevails over that which arises from the constant and inseparable union of causes and effects. Here we must not be contented with saying that the vividness of the idea produces the belief: We must maintain that they are individually the same. (T 116)

For example, even though my natural experience may enforce the belief that the earth is flat, the idea that it is in fact round is even more vivid than the first idea due to the connections fostered in me by the instruction and behavior of others. In like manner, as I have argued, moral education can produce moral beliefs. The problem in this account for the purposes of the present discussion, however, is that it is not clear that *any* beliefs for Hume produce or constitute motives (i.e., certain passions).

In fact, it is common to understand Hume as maintaining that all beliefs are products of reason, and reason alone cannot motivate. Therefore, no beliefs can motivate by themselves and beliefs do not generate motives on their own. However, one might argue, as Annette Baier does,²⁴ that this standard reading of Hume is wrong in light of what he says in the section in Book I on "Of the influence of belief" (T I 3 x): "Wherever we can make an idea approach the impressions in force and vivacity, it will likewise imitate them in its

influence on the mind..." (T 119). Impressions of sensation can "influence the mind" in producing beliefs, but by themselves do not generate motives. Hume classifies impressions of reflection, however, as passions, the class of perceptions from which motives come (although not all passions are motives²⁵); impressions of reflection are those perceptions which arise from reflecting upon the objects which are sources of pleasure and pain (T 8). So, if there are grounds to think that moral beliefs inculcated by instruction and exhortation have the effects of impressions of reflection, then these beliefs may motivate. In that case it would not be a mystery how people are generally motivated by morality even though they do not experience the natural sentiments to which the moral judgments correspond: moral education is the obvious answer.²⁶ I want to argue, however, that this answer cannot be Hume's.

How does moral education work to influence our behavior—or at least our motivations, even though those motivations may not be strong enough to produce the desired actions? Children are taught to respect the difference between right and wrong by physical and psychological reward and punishment, including verbal reinforcement—preaching, lecturing, praising, and blaming, etc. In Humean terms the goal is to foster an association between the ideas of virtue and of pleasure on the one hand, and the ideas of vice and of pain on the other, so that the thought of virtue produces a passion of attraction and the thought of vice a passion of aversion (those passions being impressions of reflection). One requirement for success in this moral training is that the initial connection that forms between the idea of virtue and the ideas of specific rewards that are pleasurable be replaced by a natural association between the former idea and the idea of pleasure itself; likewise, the association nurtured between the idea of vice and the ideas of specific punishments that are painful must be replaced by a natural association between the ideas of vice and of pain itself. Otherwise, the motivations being nurtured are motives that have nothing to do with *morality*, but are best described as fear of harmful consequences, desire for the good opinion of others, concern that one may be ostracized, and so on. In that case, when these incentives are removed, the agent loses the motives conducive to virtuous behavior because she has not acquired genuine motives to virtue.

I see nothing, however, in Hume's moral psychology that explains how the association between, say, the ideas of vice and of punishment, is somehow replaced by an association between the ideas of vice and of pain. Rather than seeing the effect of moral education as that of producing moral sentiments, it makes sense to say that moral education proceeds by working to strengthen and direct natural sentiments for morality that agents already possess to some degree or another. This conclusion is in fact consistent with the implication of Hume's famous remark, "If morality had naturally no influence on human passions and actions, 'twere in vain to take such pains to inculcate it; and nothing would be more fruitless than that multitude of rules and precepts,

with which all moralists abound" (T 457). Here Hume is suggesting that moral instruction is useful only because morality *does* initially affect our natural motivations; our natural motivations are not implanted *by* moral exhortation.

IV. Motivating Sentiments and Moral Judgments

Does this discussion indicate, then, that we can learn to *make* moral distinctions without the appropriate natural sentiments, but that we cannot learn to be *motivated* by them without the prior natural sentiments? Hume notes that even the ability to grasp moral concepts requires the possession of certain sentiments: A very kind person cannot understand what cruelty is; a very selfish person cannot grasp the notion of generosity (E 20). (Analogously, a person who has never heard a note in her life cannot really understand a Mozart piece.) Therefore, *both* the cognitive and the conative features of the moral sense presuppose that the observer/agent possess certain mental traits or be disposed in certain ways.²⁷ Is the difference between the cognitive aspect of the moral sense and the conative aspect then that we typically meet the prerequisites for grasping moral concepts (for moral cognition), whereas we typically do not meet the prerequisites for moral motivation (for moral conation)? Since we generally lack the psychological ability to divest ourselves of our particular perspectives, does this mean that we fail to fulfill the conditions necessary to moral motivation?

At last I want to suggest that there is a plausible way to understand how within Hume's theory the conditions for moral motivation are met in typical cases after all. This interpretation shows the conditions to be analogous to, and as frequently obtained as, the prerequisites for moral cognition. The suggestion is this: Just as one only need be naturally benevolent to a degree in order to understand what it is to be a benevolent person, so too, having sentiments that *approach* those that would be experienced in a GPV and *finding those* sentiments motivating may be sufficient to consider one to be motivated by moral considerations. While, for instance, I do not feel the intensity of disapproval that would correspond to the vice of the malice I find in someone I'm fond of, I do feel *some degree* of disapproval, and that may be sufficient to inspire the passion necessary to say I have a motive not to be the same way. Likewise, I may, by sympathizing with the recipients, experience approval at the thought of someone's donating money to provide a number of homeless people with an evening meal even though I cannot duplicate the recipients' actual feelings. The approval may nevertheless provoke a motivating passion toward some type of charitable behavior. Using my understanding of the GPV and my natural sentiments, I approximate to the proper feeling, even if I do not actually experience it; on this proposal, the approximation is sufficient to render the correct judgment, and the degree of approval or disapproval I in

fact feel is sufficient to provide the purported motivation.²⁸

Being excited by morality, on this account, consists in feeling an attraction to virtue via the sentiments from which we extrapolate to our moral judgments. That is, both the judgments and the motivation are derived from the same feelings—our own feelings—in some way. However, the judgments do not correspond *directly* to the feelings we in fact have; they correspond to the feelings we would have in the GPV. Is this plausible as a reading of Hume or is it simply an *ad hoc* solution to a puzzle?

I think it is a plausible interpretation. One implication of this reading is that either the sentiments that motivate us when we make moral distinctions are not *moral* sentiments (since they do not correspond directly to the moral judgments), or the moral *sentiments* simply do not track the corrected moral *judgments* precisely. I find the latter implication unproblematic: it simply implies that taking up the GPV is a matter of degree. Our sentiments still count as moral sentiments as long as they approximate to those that would be felt by one who does sympathize, unaffected by personal connections, with the narrow circle of an agent. How closely they must approximate to such feelings is difficult to say, but that Hume describes a faculty—imagination—to show how we move our feelings toward a more general perspective makes it reasonable to think he sees the GPV as a matter of degree. One arrives at proper moral judgments while being affectively engaged by sentiments that approximate to those from which the moral judgments are projected. The sentiments that engage one are still moral motives even though they do not correspond exactly to the judgments. My disapproval of a dictator in a distant country may not be as intense as my disapproval of a local politician, but I can be motivated to take action to stop the dictator's actions (e.g., lobbying my government to impose economic sanctions against the dictator), or to avoid developing the sort of dispositions that would make me insensitive to others' suffering in the way he is.

On my interpretation, then, it will not do to attribute moral motivation straightforwardly to the effects of moral education, since no amount of exhortation and/or reward and punishment will *originate* sentiments for virtue. Moral education still has a crucial role on this interpretation of Hume's view, however: to make effective in action the natural sentiments we already possess. It is one thing to have a motive in favor of morality and another thing to act on it. When Hume argues that what we call justice is conventional but that our sense of its morality is due to our natural sentiments, he concludes:

Tho' this progress of sentiments be *natural*, and even necessary, 'tis certain, that it is here forwarded by the artifice of politicians, who, in order to govern men more easily, and preserve peace in human society, have endeavored to produce an esteem for justice, and an abhorrence of injustice. This, no doubt, must have its effect; but

nothing can be more evident, than that the matter has been carry'd too far by certain writers on morals, who seem to have employ'd their utmost efforts to extirpate all sense of virtue from among mankind....The utmost politicians can perform, is, to extend the natural sentiments beyond their original bounds; but nature must furnish the material, and give us some notion of moral distinctions. (T 500)

I have shown how moral motivation in Hume's theory can still *begin* with our natural sentiments, despite our psychological inability to shed our particular perspectives completely. We have taken up the general point of view—the *moral* point of view—when we start from our real reactions to agents' characters, broaden our sympathies as far as they will go, project then how we would feel if we could sympathize with the narrow circle of the agent, and correct our judgments by that projection. This allows us to say that our moral judgments are *based on* our motivating moral sentiments.

REFERENCES

Various drafts of this paper have been presented to: "Hume in Cincinnati," the 27th Annual Philosophy Colloquium, University of Cincinnati, May 1991; the Department of Philosophy, University of North Carolina at Chapel Hill, September 1991; and the Nineteenth International Hume Conference, University of Nantes, Nantes, France, June–July 1992. I thank members of those audiences for their remarks, especially Simon Blackburn, Ted Morris, Geoffrey Sayre-McCord, and Corliss Swain. I thank my commentator at Nantes, Stanley Vodraska, for his response. I am grateful to Michael Gill, Gerald Postema, and Carol J. White for very helpful discussions of this material. For their written comments on earlier drafts of this paper, I thank Charlotte Brown, Sherryl Kuhlman, Wade Robison, Nicholas Sturgeon, and two anonymous referees for *Hume Studies*. Work on this paper was supported by a 1991 Paul Locatelli Junior Faculty Fellowship from Santa Clara University.

1 Stephen Darwall, *Impartial Reason* (Ithaca: Cornell University Press, 1983), 53 note 4.

2 Philippa Foot, "Hume on Moral Judgment," *Virtue and Vices and Other Essays in Moral Philosophy* (Berkeley and Los Angeles: University of California Press, 1978), 79. Darwall also holds in *Impartial Reason* that moral sentiments motivate directly for Hume, but has recently changed his mind; see "Motive and Obligation in Hume's Ethics," *Noûs* 27 (December 1993): 415–48.

3 More specifically, Mackie has attributed to Hume a projectivist view, which is an emotivism coupled with an "error" theory. In its simplest description, emotivism is the view that moral judgments consist in feelings, and the error theory explains why we mistakenly think our feelings represent facts external to us: we project our sentiments onto the extramental world. See J. L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth, Middlesex,

England: Penguin Books, 1977). See also: Mackie, *Hume's Moral Theory* (London: Routledge & Kegan Paul, 1980), 73–75. Simon Blackburn implies that Hume is an emotivist of a sort in his book *Spreading the Word* (Oxford: Clarendon Press, 1984); furthermore, he allies Hume with an expressivist position in "Hume on the Mezzanine Level," *Hume Studies* 19 (November 1993): 273–88.

4 See Gilbert Harman, "Moral Agent and Impartial Spectator," *The Lindley Lecture* (University of Kansas, 1986); John Rawls, *A Theory of Justice* (Harvard University Press, 1971), 183–92; and David A. J. Richards, *A Theory of Reasons for Actions* (Oxford: Clarendon Press, 1971), 86–87.

5 Harman, 2.

6 Rawls, 184.

7 It is not clear what constitutes these contradictions Hume is referring to. First, they might result from discrepancies among the way the same observer feels toward the same situation at different times, or they may result from differences in feeling among various observers toward the same situation. Second, no matter which variations Hume is referring to, it also isn't clear how these differences constitute contradictions. They produce contradictions when these observers assume that their feelings represent something beyond themselves, but what spectators might think their feelings represent before they take up the general point of view is a question I won't attempt to answer here. It seems these cannot be conflicts about morality itself, since we have yet to take up the point of view in which morality is derived. For an analysis of why Hume thinks we assume a general point of view to make our moral distinctions, see Geoffrey Sayre-McCord, "On Why Hume's General Point Isn't Ideal—and Shouldn't Be," *Social Philosophy and Policy* 11 (January 1994).

8 David Hume, *A Treatise of Human Nature*, edited by L. A. Selby-Bigge, 2nd ed. revised by P. H. Niddich (Oxford: Clarendon Press, 1978); hereafter references cited in the text.

9 This interpretation of the argument from moral motivation against the rationalists is not undisputed. My reading here commits Hume to "moral internalism," the view that moral judgments (the making of the moral distinctions, which for Hume consists in having feelings) motivate directly, on their own. Some philosophers have argued contrary to this interpretation that it is Hume's view that "external" sanctions, such as fear of punishment or concern with good reputation, provide the incentives to morality. I present a full and detailed defence of the internalist interpretation in my paper "How Does the Humean Sense of Duty Motivate?" presented to the 20th International Hume Conference, University of Ottawa, Ottawa, Canada, July 1993. However, for the purposes of this discussion, it only need be apparent that there is ample evidence for the internalist reading, so that it is obvious that those who subscribe to the sentimentalist reading of Hume have good reason to do so.

10 I thank an anonymous referee for *Hume Studies* for raising this question.

11 I have not offered a comprehensive interpretation of the difficult "wilful murder" passage, which would require detailed argument and take me far

afield of the focus of this paper. However, I want to note that I am greatly influenced by Nicholas Sturgeon's reading of this passage in his unpublished paper, "Moral Skepticism and Moral Naturalism in Hume's *Treatise*," in which he argues that this text and others show that Hume believes there are internal (moral) matters of fact discovered by feeling as opposed to external (non-moral) matters of fact discerned by reason.

12 Geoffrey Sayre-McCord, in his unpublished paper "Hume and the Bauhaus Theory of Ethics," argues that Hume's talk of multiple general perspectives is significant, since, on Hume's view, we approve of what we find agreeable and useful from a mutually accessible point of view, and this, according to Sayre-McCord, may not necessarily be the same for everyone or the same throughout time.

13 This implies that we overlook the momentary appearances of things, our present situation, and our own interests; we do not form judgments of the characters of persons "merely from the tendency of their characters to our own benefit" (T 582–83). This is not to say that Hume describes the GPV as an antidote to self-interest, since he doesn't view us as basically egoistic; rather, it is just another way of saying that the GPV counteracts the effects on our sympathies of our particular relations to others.

14 I put "disinterested" here in quotes to advise caution about how the term should be read. I don't intend it here to mean that we distance ourselves from others' concerns when we take up Hume's GPV, but that we distance ourselves from our own particular concerns. I want none of the shorthand ways in which I refer to the GPV to imply that Hume's theory is to be read as an "ideal observer" theory, since I reject that reading for reasons which will become clear later.

15 David Hume, *An Enquiry Concerning the Principles of Morals*, edited by L. A. Selby-Bigge, 3rd ed. revised by P. H. Nidditch (Oxford: Clarendon Press, 1975; hereafter references are cited within the text.

16 Hume notes as well in his essay "Of the Standard of Taste" that people very rarely experience the feelings indicative of the generally accepted standards for beauty and fine taste. Those sentiments, Hume explains, are "of a very delicate nature, and require the concurrence of many favorable circumstances to make them play with facility and exactness, according to their general and established principles" (David Hume, *Essays: Moral, Political, and Literary*, edited by Eugene F. Miller (Indianapolis: Liberty Classics, 1987), 232). That our distinctions of beauty do not correspond directly with the reactions of most individuals is not as problematic for a theory of taste as it is for a theory of morality in at least one respect: We do not expect distinctions of beauty to be directly connected to action in the way that we expect distinctions of virtue to be. I do not intend my remarks here to indicate necessarily that Hume's spectator in matters of taste has the same qualities as his spectator in matters of morality, however; this is a topic for another discussion.

17 Carole Stewart ("The Moral Point of View," *Philosophy* 51 [April 1976]: 177–87) maintains that the requirement that we take up the GPV gives reason no more significant causal role in the formation of moral beliefs on Hume's account than that of causing us to reconsider the situation. I think such a

response to the question about the role of inference in our making moral distinctions is inadequate, since our reconsidering a moral judgment requires a method by which we change the judgment, and that method is not perceptual. On reconsideration, our perception will still be the same, but the judgment will be different.

18 In "Morality as a Back-Up System: Hume's View?" *Hume Studies* 14 (April 1988): 25–52, Marcia Baron notices a tension related to this concern when she asks how the corrected sentiments are "natural" in the way that a person's natural, unregulated passions are. She doesn't, however, make anything out of the claim that it is actually our judgments that we correct, rather than the sentiments themselves.

19 Hume's contemporary, Francis Hutcheson, offers a moral sense theory that doesn't raise this question. Hutcheson realizes that the sentiments we in fact experience cannot always be considered a source for morality and must be "corrected" in the way we correct our perceptions of colors when we see objects under non-standard circumstances. See Hutcheson, *Illustrations on the Moral Sense* (1728), Section IV, in *An Essay on the Nature and Conduct of the Passions and Affections with Illustrations on the Moral Sense* (Gainesville, Florida: Scholars' Facsimiles & Reprints, 1969), 286. He holds that we are motivated by desires and affections, but does not maintain that our moral distinctions motivate; consequently, that our moral judgments sometimes, or even often, require inference beyond our own feelings poses no problem for his theory of motivation.

20 Adrian Piper solves the problem of taking up the GPV by arguing, on the basis of evidence from the *Enquiry*, that it is really the stance of enlightened self-interest, as opposed to what she calls the stance of "strict impartiality" ("Hume on Rational Final Ends," *Philosophy Research Archives* 14 [1988–89] 193–228). The perspective of enlightened self-interest is that viewpoint I assume when I put aside concerns peculiar to my present situation and consider instead the concerns common to myself over time, including those interests I'll have in the future.

There are, however, no compelling grounds to think the GPV is to be understood as enlightened self-interest. In the *Enquiry*, Hume argues against the theories that attempt to explain our approval of the social virtues (the virtues of justice and benevolence) entirely in terms of self-interest (see especially E 214–20). Hume clearly holds there that when we take up the GPV, we approve of persons whose character traits are conducive to the happiness of those whose lives are touched by those persons. If the GPV were the stance of enlightened self-interest, Hume would be committed to the implausible claim, which at one point he explicitly argues against, that *my* long-term self-interest is served when the interests of other persons—even persons spatially and temporally remote from me—are served (see E 215–16). So, while reinterpreting the GPV in terms of long-term self-interest suggests a way to have our actual sentiments conform to our judgments, it is plainly not what Hume had in mind.

21 For a discussion of the pervasiveness of the use of general rules in all aspects of Hume's philosophy, see Thomas K. Hearn, Jr., "General Rules in Hume's *Treatise*," *Journal of the History of Philosophy* 8 (1970): 405-22.

22 If the perceiver did experience the pleasures and pains that directly correspond to her moral distinctions, an account of how those distinctions motivate can fairly easily be constructed from the materials Hume offers: The feelings of pleasure and pain (of which moral approbation and disapprobation are species) evoke (or perhaps are even identified with) all sorts of motivating passions, including "with the least preparation...the direct passions of desire and aversion, grief and joy, hope and fear, along with volition" (T 438). Alternately, moral approbation produces—or may be identified with—the indirect passions of love and pride, and moral disapprobation the indirect passions of hatred and humility, the former of each pair being connected to a motivating passion (see, e.g., T 574). The indirect passions arise when the objects that cause pleasure or pain acquire a relation to ourselves or others (see, e.g., T 277-79, 229-30, 574). However, it must be possible for Hume to explain how moral pronouncements that are the result of "correcting the language" in accord with the GPV but which don't correspond to our actual sentiments are somehow linked to or identified with motivating passions.

23 I have noted that Hume's familiar argument against the rationalists relies on the premise that morality motivates: since reason alone does not, morality cannot be a matter of reason alone. Since the rationalists against whom Hume is arguing agree that morality always derives from reason by itself, Hume only needs to show that in at least one case it does not; so, one may contend, the success of his argument requires only a commitment to the claim that in at least one instance morality motivates. It is erroneous, however, to use the narrow context of this argument as evidence of the strength of Hume's claim; it is obvious from other contexts that he thinks his observation is widely substantiated. (See especially T 455-56.) It trivializes an important component of Hume's moral philosophy to argue that his observation about the practicality of morality amounts to the view that morality imparts a motive, but only rarely.

24 Annette Baier, *A Progress of Sentiments* (Cambridge: Harvard University Press, 1991), 152-73. Baier rejects the traditional reading of Hume that reason cannot originate motives in favor of an interpretation of T I 3 x that has beliefs which are products of reason producing motivating passions. In the section, "Of the influence of belief," Hume claims, for example, "The effect of belief... is to raise up a simple idea to an equality with our impressions, and bestow on it a like influence on the passions" and "[a] belief is almost requisite to the exciting our passions, so the passions in their turn are very favorable to belief..." (T 119-20). I believe, however, that there is an interpretation of such claims that can reconcile them with the standard reading of Hume on the inefficacy of reason to motivate by itself, and so my discussion in the text of this paper supposes the standard line. I cannot develop my argument here, however.

25 It is not clear what criteria determine what passions are motives for Hume, but it makes it clear that some are not: Pride and humility are not motives, and

love and hatred do not themselves motivate, although these two are connected respectively to the motives of benevolence and anger (T 367).

26 This answer has been suggested to me in conversation by numerous people. Stephen Darwall takes this line when he argues that the motivation to be virtuous comes not from moral sentiments but from inducements such as fear of punishment, pressure from political oration, and concern for one's reputation. See Darwall, "Motive and Obligation in Hume's Ethics."

27 This is not to say that one would need to have certain virtues to understand how people use moral terms. One may learn by verbal reinforcement how people talk about morality and immorality—for example, that they call revengeful behavior "vicious"—in a way analogous to how a sightless person can learn how sighted people talk about what they see. The result of such instruction would be to teach people (most likely children) about a matter of fact in the external world, the sort of fact that can also be grasped by cause-and-effect reasoning, namely a fact about what people consider virtuous and vicious. This is obviously not the sort of acquisition of moral concepts that Hume has in mind when he says that very kind people cannot really understand gross selfishness.

28 This proposal was first suggested to me in discussion by Corliss Swain and Ted Morris. I attempt to develop the idea in some detail here.

Received June 1993

Revised February 1994