



PROJECT MUSE®

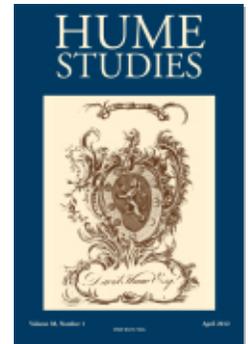
---

## Humean Sentimentalism and Non-Consequentialist Moral Thinking

Michael B. Gill

Hume Studies, Volume 37, Number 2, November 2011, pp. 165-188 (Article)

Published by Hume Society  
DOI: [10.1353/hms.2011.0656](https://doi.org/10.1353/hms.2011.0656)



➔ For additional information about this article

<http://muse.jhu.edu/journals/hms/summary/v037/37.2.gill.html>

# Humean Sentimentalism and Non-Consequentialist Moral Thinking

MICHAEL B. GILL

*Abstract:* Of the many objections rationalists have raised against moral sentimentalism, none has been more long-lived and central than the claim that sentimentalism cannot accommodate the non-consequentialist aspects of our moral thinking. I examine how Stephen Darwall directs this criticism at Hume's account of moral judgment and argue that Darwall's criticism is based on an incorrect interpretation of Hume's view of motivation and the moral sentiments. Humean moral psychology is more nuanced than Darwall's objection in particular and rationalist criticisms more generally have assumed. Developing a clear picture of why Hume's account of moral judgment does not imply an implausible consequentialism reveals the strength of Hume's moral sentimentalism overall.

## 1. The Non-Consequentialist Objection

Of the many objections moral rationalists have raised against moral sentimentalism, none has been more long-lived and central than the claim that sentimentalism cannot accommodate the non-consequentialist aspects of our moral thinking. John Balguy raised an early version of the non-consequentialist objection just two years after Francis Hutcheson published the first systematic development of moral sentimentalism. As Balguy understood it, Hutcheson's sentimentalism implied that what makes an action virtuous is its effects, such as the advantages or pleasures it produces. According to Balguy, however, what we actually think makes

an action virtuous is an intrinsic quality it possesses, which is independent of any effects the action may have. As Balguy rhetorically asks, “Is virtue no otherwise good or amiable, than as it conduces to public or private advantage? Is there no absolute goodness in it? Are all its perfections *relative* and instrumental?”<sup>1</sup> Actions produce consequences, and some of those consequences may be “natural goods.” But the morality of an action is independent of the natural goodness it brings into existence. As Balguy puts it, “*moral good* is an *end*, an *ultimate end* of one kind, as *natural good* is of another. . . . However pleasure may be the consequence or appendage of virtue, yet, strictly speaking, it is not the end of a moral agent, nor the object of a moral affection, but virtue alone, antecedent to all considerations, and abstracted from every natural good.”<sup>2</sup>

Kant objected to sentimentalism on similar grounds. According to Kant, sentimentalism can explain only the motives we have to promote “a certain interest [whether it be] one’s own or another’s;” in other words, sentimentalism can account only for the ends that a person has that consist of “*what he intends to accomplish* through his action.”<sup>3</sup> The ends of morality, however, are “abstract[ed] completely from every end that has to be *brought about*” or “produced.”<sup>4</sup> The goal of producing desirable states of affairs, which is the only kind of goal sentimentalism can accommodate, can give rise only to hypothetical imperatives, but morality “can be expressed only in categorical imperatives.”<sup>5</sup>

Like Balguy and Kant, Rawls thought that sentimentalism—or at least *Hume’s* sentimentalism—was unable to explain the moral concerns we have that are not based on the desirability of outcomes. In explaining this idea, Rawls distinguished between “object-dependent desires” and “principle-dependent desires.”<sup>6</sup> An object-dependent desire is a desire to produce certain states of affairs. A principle-dependent desire is a desire to act on a principle distinct from the states of affairs that acting in that way will produce. According to Rawls, Hume’s sentimental psychology explains only object-dependent desires—only desires to produce certain results.<sup>7</sup> But in fact we also have principle-dependent desires. We sometimes prefer a course of action even though its outcomes are less desirable; we sometimes perform actions because we think they are required by a principle that is not results-oriented. And, according to the objection, Hume’s sentimentalism cannot accommodate this.

The most important recent version of this non-consequentialist objection to moral sentimentalism is Stephen Darwall’s.<sup>8</sup> Darwall’s objection is aimed not at sentimentalism in general (he thinks that Adam Smith’s sentimentalist view is not susceptible in the same way<sup>9</sup>) but specifically at Hume’s account.

In what follows, I will argue that Darwall’s objection is based on an incorrect interpretation of Hume’s account of motivation and the moral sentiments. Humean psychology has a more nuanced and multifarious story to tell about these things than Darwall’s objection and other rationalist criticisms have assumed.

Developing a clear picture of why Hume's approval-based account of moral judgment does not imply an implausible consequentialism will, thus, help to reveal the strength of Hume's moral sentimentalism more generally.

According to Darwall, Hume's approval-based account of morality can accommodate only judgments based on desires for states of affairs.<sup>10</sup> As a result, according to Darwall, Humeanism cannot account for commonsensical opposition to certain kinds of moral trade-offs. As Darwall puts it,

In accepting a norm of justice requiring me to restore a seditious bigot's property, I take there to be a reason for doing so that cannot be reduced to a reason for (desiring) the existence of any state, even the state of the money's being returned. Perhaps the world would be a better place if the money went to Oxfam; that might be a more desirable state of affairs. And even if I think that a property-restoring acts' being done is a better state, the reason I will credit in accepting a norm of justice requiring me to restore the property will differ from any deriving from (or consisting in) the value of that state. Imagine . . . that I can bring about more such valuable states by the shocking spectacle of violating the norm myself thereby causing, say, two other would-be violators not to go through with their previously intended violations. The (agent-relative, second-person) reasons for acting that derive from a norm of property would not recommend that I do so. Reasons for action cannot, in general, be reduced to reasons to desire states, and, in particular, a reason of justice to return a seditious bigot's property cannot be. Or so someone who has the motive of justice must think if she is to regulate her conduct in the way she must for a whole plan or scheme to be collectively beneficial. But . . . this conflicts with Hume's general theory of motivation.<sup>11</sup>

As Darwall has it, in returning the money to the seditious bigot, the just person is not trying to produce a certain outcome. She is not trying to bring the world into accord with her desire for a certain state of affairs. She is, rather, acting on principle, heeding the normative call for an action "pure and simple."<sup>12</sup> This normative call is impervious to the lure of trading one unjust action (of one's own) for two just actions (of others). That's because the attraction of such a lure would be based on the desirability of certain outcomes, but the normative call (as Darwall sees it) is not based on the desirability of certain outcomes. It's a call to follow a norm, not a desire that has a state of affairs as its object. According to Darwall, however, the official Humean view takes all reasons to be based on desires that do have outcomes or states of affairs as their objects. More specifically, the Humean view takes all of an agent's reasons to be based on the agent's desire to experience more rather than less pleasure. As Darwall puts it in an earlier work, "Hume's theory of action thus

not only employs the traditional idea that the will invariably aims at the good; it also interprets that idea hedonistically and egoistically.”<sup>13</sup> And if the only reason a person has to perform an action is that it will produce an outcome whose observation will give her a pleasurable feeling (of, say, approval), then she will have an even more powerful reason to perform an action that produces an outcome whose observation will give her two pleasurable feelings (of approval). Because the Humean agent will have every reason to trade off something that will cause her to feel one pleasure (of approval) for something that will cause her to feel two pleasures (of approval), Hume cannot accommodate a reason not to trade off one of one’s own just acts for two just acts of other people. But we *do* think there is a reason not to trade off one of one’s own just acts for two just acts of other people. The Humean view thus fails as an account of our moral thinking.<sup>14</sup>

In response to this objection, I will try to show that the Humean theory can in fact accommodate our non-trade-off judgments and is thus not open to the non-consequentialist objection. The approach I will take to defending Hume is not the only possible one. We could also highlight Hume’s acknowledgement that some of our motives (such as to punish our enemies, benefit our friends, and be kind to children) are unconnected to any thought of producing pain or pleasure; he writes, “Beside good and evil, or in other words, pain and pleasure, the direct passions frequently arise from a natural impulse or instinct, which is perfectly unaccountable” (T 2.3.9.8). Or we could follow Garrett, who has defended Hume by explaining his capacity to accommodate the motive to follow a “policy” that has long-term benefits rather than simply to act on a “case-by-case evaluation of interest,” which would explain why a person may in a particular situation perform an action that she does not expect to produce an outcome that will give her pleasure. But what I will try to show is that Humean moral psychology includes pleasures that accord perfectly well with the non-trade-off moral judgments at the heart of the objection, and that Darwall and others who have advanced the objection have failed to notice this because they have attributed to Hume a view of pleasure that is cruder than that which he actually develops.

## 2. Beauty and Trade-offs

The non-consequentialist objection moves from the claim that Humean sentimentalism can account only for desires for states of affairs whose observation one finds pleasurable to the claim that Humean sentimentalism cannot account for our opposition to certain kinds of moral trade-offs (such as our opposition to performing one unjust action in order to cause other people to perform a multitude of other just actions). But there is a desire for a type of pleasurable state of affairs that would block that move: namely, an agent’s desire for the state of affairs in which *she herself* not perform an unjust act. I may desire that the world contain as

many just acts as possible, regardless of who performs them. But I may also have a distinct desire that the world contain just acts *performed by me*. That second desire may not be merely another instance of the first. And the pain I receive from performing an unjust act myself might be so much weightier than the pain I receive from observing two others' unjust acts that we do not need to refer to anything else to explain my resistance to trading off one of my just acts for two of two others.

The distinction between agent-relativity and agent-neutrality is helpful in explaining this point. An agent-neutral desire has as its object a state of affairs that can be described without any reference to the agent who has the desire. An agent-relative desire has as its object a state of affairs that cannot be described without reference to the agent who has the desire. The agent-neutral desire for a world in which just actions are performed may not be able to fund the non-trade-off judgments at the core of the non-consequentialist objection. But the agent-relative desire for a world in which I myself perform just actions can fund such judgments. So if Humeanism can accommodate agent-relative desires of this kind, the crucial final move of the objection will be blocked.<sup>15</sup>

To make this general sentimentalist approach plausible, let's first consider artistic motivation. An artist, call her Jane, has a desire to produce a beautiful object herself, say a painting. Jane also has a desire that the world contain as many beautiful objects as possible, regardless of who produces them. From her possessing those two desires, can you infer that Jane will prefer not to create a painting she wants to create if she believes that by doing so she can spur two other people to produce two beautiful paintings? No. The satisfaction Jane receives from producing a beautiful painting herself is quite different from, and quite possibly much weightier than, the pleasure she receives from observing a beautiful painting produced by someone else. The satisfaction of creating a beautiful work of art isn't simply one unit of something that the pleasure of observing two other works of art is two units of. That both feelings are pleasurable reactions to states of affairs does not give you any reason to hold that Jane will think it better to trade her own artistic creation for the observation of two others' creations. The same point can be made if we consider the desire to write a novel or run a marathon and the pleasure of reading others' novels or watching others finish marathons. You may want to write a novel or run a marathon because you think you will get great satisfaction out of it. You may also want two other people to finish novels or marathons because you think you will enjoy reading or watching them. I cannot, however, infer on this basis alone that you will think yourself better off abandoning your own novel or marathon in order to spur two other people to finish theirs.

It is not even the case that the pleasure Jane receives from observing an already-existent beautiful object will necessarily have more deliberative weight with her than the pleasure she would receive from two new beautiful objects' coming into existence. Let us say that Jane observes an object and that her observation is

pleasurable in a way that leads her to judge it beautiful. Does this give you reason to think that she will, all other things being equal, choose to destroy the original object so that two other objects that give her the same kind of pleasure will be brought into existence? I think it's pretty clear that it does not. Imagine that the original object is a painting. Jane realizes that by destroying the painting she will inspire multiple other people, spurred to artistic creation by fury at her destructive act, to create paintings that she will find just as beautiful as the original. And let us say that she chooses not to destroy the painting. Do we need to postulate that she has some desire that *conflicts* with the aesthetic pleasure she receives from the original painting in order to explain her not destroying that painting? No. Jane's aesthetic response to the original painting itself seems perfectly compatible with her choosing that the painting continue to exist even if destroying it would lead to the production of two other beautiful paintings. (This is a case I will come back to in section 4.)

But if you know that Jane finds an object beautiful, can't you infer that, all other things being equal, she will prefer that more objects like it exist rather than fewer? Perhaps. That inference, however, does not imply any consequentialist-vs-non-consequentialist distinction between the pleasure Jane receives from a beautiful object and her reasons for moral action. From the fact that Jane thinks she morally ought to perform a particular action it may be legitimate to infer that, all other things being equal, she will prefer more actions like it be performed rather than fewer. But what the non-consequentialist objection requires is that her receiving pleasure from an object's having a certain feature warrants the inference that she will choose a state of affairs in which that original object is destroyed but more objects like it exist over a state of affairs in which the original object persists but fewer objects like it exist overall—and this inference is not warranted. Concluding that Jane will prefer the trade-off of one actual desirable object for the bringing into existence of multiple possible desirable objects is just as unwarranted with regard to her responses to beautiful objects as it is with regard to her responses to acts of justice. Her aesthetic appreciation of a beautiful object gives you no more reason to think that she would destroy that object in order to spur other people to create multiple beautiful objects than her appreciation of just action gives you reason to think that she would act unjustly in order to spur other people to act justly.

How does this discussion of Jane's aesthetic reactions help the Humean sentimentalist respond to the non-consequentialist objection? The aesthetic case helps because it is an example of a set of evaluations and motivations that are plausibly taken to be based on sentiment and yet do not lead to the intrapersonal trade-offs at the heart of the objection. If we find nothing amiss in combining a sentimentalist account of aesthetics with non-consequentialist features of our aesthetic reasons, then why think there is anything amiss in combining a sentimentalist account of morality with non-consequentialist features of our moral reasons?<sup>16</sup>

The satisfaction Jane receives from creating a beautiful work of art is not merely one unit of a kind of pleasure that she receives two units of when she observes two other artists' works of art. There is thus no conceptual bar to the idea that the pleasure that motivates Jane to perform a just act herself is not merely one unit of the pleasure she could receive two units of by observing two other people perform just acts. In the next four sections, I will try to show that the details of Hume's account of pleasure, motivation, and evaluation cohere perfectly with this idea.

### 3. The Non-fungibility of Humean Pleasures

Bentham thought that all pleasures were fungible and that as a result decisions about what to do were entirely quantitative matters.<sup>17</sup> This Benthamite position is based on the idea that all pleasures can be reduced to the selfsame impression, to a single, distinct psychological state.

Hume rejects this view of pleasures as fungible. Humean pleasures do not all reduce to quantifiable, commensurable units, for Hume denies that there is a simple impression of pleasure *per se*. He writes:

[T]is evident, that under the term *pleasure*, we comprehend sensations, which are very different from each other, and which have only such a distinct resemblance, as is requisite to make them be express'd by the same abstract term. A good composition of music and a bottle of good wine equally produce pleasure; and what is more, their goodness is determin'd merely by the pleasure. But shall we say upon that account, that the wine is harmonious, or the music of a good flavour? In like manner an inanimate object, and the character or sentiments of any person may, both of them, give satisfaction; but as the satisfaction is different, this keeps our sentiments concerning them from being confounded, and makes us ascribe virtue to the one, and not to the other. Nor is every sentiment of pleasure or pain, which arises from characters and actions, of that *peculiar* kind, which makes us praise or condemn. (T 3.1.2.4; SBN 472)<sup>18</sup>

While there are multiple impressions that are pleasurable, pleasure simpliciter is, for Hume, merely an abstract idea. We can form a mental grouping or associated set of all impressions that are pleasurable, but there is no one impression that is simply pleasure itself.

This feature of Hume's account should immediately make us question whether he is committed to holding that a person will always prefer two things that give pleasure to one thing that gives pleasure. For Humean pleasure is not a single, discrete mental experience that can be chopped up into fungible units that can then be compared in the way that consequentialist trade-offs require.

My appreciation of a musical composition and my appreciation of a fine bottle of wine are both determined by the pleasure they give me. But that does not imply that I would trade the single pleasure of listening to one musical composition for the two pleasures of drinking two bottles of wine. The pleasures involved aren't sensibly chopped up, quantified, and measured in the way that such trade-offs presuppose.

Someone might try to reassert the objection by claiming that it requires only trade-offs between experiences of approval. The charge is that Hume is committed to holding that we will always prefer two experiences of approval (of the moral acts of others) to one experience of approval (of my own moral act). And even if trade-off implications don't follow in any clear way in cases in which there are different kinds of pleasures, they do follow in cases in which it is approval that is on both sides. Comparing the pleasure of a bottle of wine to the pleasure of a musical composition might be akin to comparing apples to oranges. But comparing approvals to approvals is apples to apples.

This reassertion of the objection fails, however, because even comparing two pleasures that go by the same name can be a case of comparing apples and oranges. The aesthetic case is once again instructive. The desire an artist has to create a beautiful painting and the desires an artist has for multiple other people to create beautiful paintings might both be said to be based on aesthetic pleasure. But that does not give us grounds for claiming that the artist will prefer to forgo the single pleasure of creating a beautiful painting so that she can experience the two pleasures of observing two other persons' beautiful creations. (Indeed, there's even something deeply suspect about this quantitative way of speaking about aesthetic pleasure.) Similarly, your receiving pleasure from a painting you find beautiful does not warrant my believing that you will prefer that painting be destroyed so that two equally beautiful paintings will be created. And Hume points out that sentiments that are classified as approval differ from each other in a similar way. The sentiment of approval that one virtue elicits can be "somewhat *different* from that, which attends the other virtues. . . . Each of the virtues, even benevolence, justice, gratitude, integrity, excites a different sentiment or feeling in the spectator" (T 3.3.4.2; SBN 607; see also E 7.29; SBN 260 and App 4.3–6 and 21–22; SBN 314–16 and 322–23).<sup>19</sup> The virtues that make us love our friends lead us to feel a sentiment toward them that differs from the sentiment that the virtues of heroes we admire make us feel. Approval of a trait agreeable to its possessor is not a smaller amount of the same thing that approval of a trait useful to others is a larger amount of. Approval of a trait that is useful to its possessor is not a smaller amount of the same thing that approval of a trait that is useful to others is a larger amount of. Hume makes all of these points in his pluralist account of the bases of our moral judgments.<sup>20</sup> All of these points fit with his view of pleasures as non-fungible, the idea that pleasurable sentiments do not all reduce to some quantity

of one selfsame impression or mental state. (Hume would say that pleasure *per se* is merely an abstract idea.) And these points about the variability of moral pleasure enable Hume to resist the implication of the simplistic trade-offs at the heart of the non-consequentialist objection—just as he can resist the idea that we would always trade one beautiful musical composition for two good bottles of wine.

Most importantly, at the heart of the non-consequentialist objection are *interpersonal* trade-offs—trade-offs of one approval-of-myself for two approvals-of-others. But Hume's account of the motivational influences of the indirect and moral passions implies that self-approval has a deliberative weight that is not comparable to the deliberative weight of approval of others. The non-consequentialist objection will work only if Hume is committed to holding that the interpersonal comparison of self-approval to approval of others is an apples to apples comparison, but on Hume's own account it is apples to oranges. The kind of asymmetry of the motivational profiles of these pleasures is something we have already seen in our discussion of artistic creation in section 2. We will now see how Hume builds that motivational asymmetry into his account of the indirect and moral passions.

#### 4. Pride and Love

Let's return to Jane. She now faces two courses of action. If she takes the first course of action, she will be proud of herself—or at least will avoid the painful feeling of humiliation. If she takes the second course of action, she will feel humiliated, but she'll also cause herself to feel love for three people toward whom she is currently indifferent. So, the first course of action promises a net result of one pleasurable impression (namely, a feeling of pride or the avoidance of a feeling of humiliation). The second course of action promises one painful impression (humiliation) and three pleasurable impressions (love for three other people), for a net result of plus two. Can we infer that Jane will prefer the second course to the first—that she'll bring on herself one feeling of humiliation so that she can also experience three feelings of love? Of course not. Pride and humiliation and love don't operate in the quantifiable, interpersonally fungible way such trade offs require. And Hume's account of the passions fits perfectly well with these non-trade-off phenomena.

On Hume's account, pride and love are pleasurable indirect passions. Neither of them is on its own an immediate motive to action (T 2.2.6.3–6; SBN 367–68),<sup>21</sup> but both influence the will in indirect ways. What I want to highlight is the difference between how the two of them influence the will.

According to Hume, my thinking that I will feel pride as a result of performing an action can never be the sole explanation of my being motivated to perform it. There must be some other consideration that initially leads me to be motivated to perform it. But my thinking that action will make me feel pride will add to that initial motive. Hume writes,

Supposing that there is an immediate impression of pain or pleasure, and that arising from an object related to ourselves or others, this . . . by concurring with certain dormant principles of the human mind, excites the new impressions of pride or humility, love or hatred. That propensity, which unites us to the object, or separates us from it, still continues to operate, but in conjunction with the indirect passions, which arise from a double relation of impressions and ideas. These indirect passions, being always agreeable or uneasy, give in their turn additional force to the direct passions, and encrease our desire and aversion to the object. Thus a suit of fine cloaths produces pleasure from their beauty; and this pleasure produces the direct passions, or the impressions of volition and desire. Again, when these cloaths are consider'd as belonging to ourself, the double relation conveys to us the sentiment of pride, which is an indirect passion; and the pleasure, which attends that passion, returns back to the direct affections, and gives new force to our desire or volition, joy or hope. (T 2.3.9.4; SBN 439)

My feeling pride in a thing cannot on its own explain my desire for it, since my feeling pride in it must itself be explained by some feature of it that I find desirable. However, the prospect of feeling pride will “encrease” or give “new force” or “additional force” to my desire or volition. When I have an antecedent desire to produce some result, my thinking that this result will cause me to feel the pleasure of pride will strengthen the antecedent desire. Perhaps in addition to wanting to buy a suit, I am also averse to spending the money that it will cost. If it were simply a contest between my desire for the suit based on its beauty and my resistance to the suit based on its cost, the latter would win out. But then I realize how proud I will feel walking into the office in my new suit, and my desire to buy it grows in strength. With the additional force that the prospect of feeling pride adds, my desire to buy the suit now wins out.

Humean pride is a feeling of pleasure (toward oneself). Humean love is a feeling of pleasure (toward another). Like pride, love is not itself an immediate motive to action. Like pride, love nonetheless influences the will in indirect ways. But the way love influences the will differs from the way pride does—and this difference accounts for Jane’s resistance to trading off one feeling of humiliation for three feelings of love.

Pride influences an agent by strengthening her motive to do something that she thinks will give her the pleasure of pride. The main influence of love, by contrast, is not to motivate an agent to bring about states of affairs in which she will feel love. Indeed, love does not influence an agent’s will chiefly by adding to, or giving rise to, motives to create for herself experiences of feelings of pleasure at

all. Love, rather, influences an agent chiefly by giving rise to a motive to make the person she loves happy. Hume writes,

The passions of love and hatred are always followed by, or rather conjoin'd with benevolence and anger. 'Tis this conjunction, which chiefly distinguishes these affections from pride and humility. . . . [L]ove and hatred are not compleated within themselves, nor rest in that emotion, which they produce, but carry the mind to something farther. Love is always follow'd by a desire of the happiness of the person belov'd, and an aversion to his misery: As hatred produces a desire of the misery and an aversion to the happiness of the person hated. . . . [B]enevolence and anger are passions different from love and hatred . . . [but] . . . conjoin'd with them, by the original constitution of the mind. As nature has given to the body certain appetites and inclinations . . . ; she has proceeded in the same manner with the mind. According as we are possess'd with love or hatred, the correspondent desire of the happiness or misery of the person, who is the object of these passions, arises in the mind, and varies with each variation of these opposite passions. (T 2.2.6.3–6; SBN 367–68)

Love influences an agent's will by causing a motive to benefit the person for whom she feels love. Love is a pleasurable sentiment, and it has motivational influence, but that influence does not mainly concern an agent's desires for states of affairs in which she herself experiences more pleasure.

Hume thinks, for instance, that we tend to feel esteem for people who have "power and riches" and to feel "contempt" for people who exhibit "poverty and meanness." He also thinks that this "esteem and contempt are to be consider'd as species of love and hatred" (T 2.1.5.1), but he does not equate our esteem for the rich with a desire for certain kinds of outcomes that produce pleasure in us. The esteem we feel for a rich person is pleasurable, but Hume nowhere suggests that such esteem is the same as or leads to a desire to observe as many rich people as possible. The state of mind of esteeming a rich person is that of having a positive feeling about her, not that of having a motive to change the world in any particular way. Just as my aesthetic appreciation of an object is distinct from my wanting to own the object or maximize the number of moments of my observation of it, so too my esteem for a rich person is distinct from my wanting anything in particular at all. Indeed, Hume takes great pains to argue that our esteem for the rich is neither equivalent to nor based on our desire for any advantages we might expect to gain from their wealth. He maintains "not only that we respect the rich and powerful, whether they shew no inclination to serve us, but also when we lie so much out of the sphere of their activity, that they cannot even be suppos'd to be endow'd with that power" (T 2.2.5.10; SBN 361). The esteem we feel for the rich

is “disinterested,” distinct from states of mind with a world-to-mind direction of fit (T 2.2.5.11; SBN 361).

Because Hume takes esteem to be a species of love, and because he claims that the passion of love “is always followed by, or rather conjoin’d with benevolence,” he must hold that when we feel esteem for a rich person, we also tend to feel a certain amount of good will toward her, even if that good will’s motivational force is relatively weak. As he explains it, “love and esteem are at the bottom the same passion” (T 2.2.2.10; SBN 337), and “benevolence attends both,” even if benevolence “is connected with love in a more eminent degree” (T 3.3.4.13; SBN 613). And it is, in fact, plausible to hold that if I esteem another person, I will, all other things being equal, prefer that her welfare be promoted rather than obstructed. That preference, however, is not necessarily an agent-neutral, maximizing desire for as many outcomes as possible that produce in me feelings of pleasure, which is what the non-consequentialist objection requires. Hume’s view may entail that the esteem I feel for a wealthy person is always conjoined with a desire for that particular person’s happiness, but it does not entail that this esteem is equivalent to, or is even always conjoined with, a preference for the state of affairs in which the original wealthy person is made poor while a few more people are made wealthy over the state of affairs in which the original person remains wealthy while no one else is made wealthy. The Humean position can consistently deny, as Hume himself surely would want to, that my good will toward a wealthy person whom I esteem implies my having the preference to destroy this person’s wealth so that different people could be made wealthy. My esteem—like my feelings of love in general—has an agent-relative character that spreads to the desires associated with them, and not the agent-neutral character that implies interpersonal trade-offs.

One way to underscore the difference between pride and love with respect to their influence on the will is to note the temporal asymmetry between them. It is the *prospect* of feeling pride that influences the will: the thought that doing something will cause me to feel pride in the future will “encrease [my] desire” to do it. But we aren’t typically motivated to perform an action because we think that it will cause us to feel love in the future for someone whom we do not currently feel love for. Rather, *occurrent* feelings of love give rise to motives to benefit the beloved in the future: we actually feel love for a person and that feeling causes us to “desire . . . the happiness of the person below’d.” Because I will feel pride toward myself by doing something, I have a stronger motive to do it. By contrast, it’s not typically the case that because I will feel love for S by doing something, I develop a motive to do it. Rather, because I occurrently feel love for S, I am motivated to do things that will make her happy.

The difference between the motivational profiles of the *occurrent* pleasurable sentiment of love and the *prospect* of feeling the pleasurable sentiment of pride in

the future helps explain what might otherwise have seemed odd about one of the aesthetic cases we discussed above. Suppose Jane observes a painting that gives her aesthetic pleasure. She realizes that if she destroys the painting, she will spur two other people to create two equally beautiful paintings, which she will then be able to observe, resulting in two aesthetic pleasures. Even if we eliminate all non-aesthetic considerations of potential punishment and property violation and the like, it is still natural to think that Jane will choose not to destroy the original painting, even if doing so will spur the creation of two equally beautiful paintings. How can Humean sentimentalists explain this? They can do so by holding, plausibly, that aesthetic pleasure is like love. Like love, aesthetic pleasure is not in and of itself a motive to action. Like love, aesthetic pleasure gives rise in Jane to a desire to promote the well-being of the thing she feels it about. In the case of love, this would be a desire for the beloved's happiness; in the case of aesthetic pleasure, this would be a desire for the preservation of the object. Crucially, her desire to promote the well-being of the object of her love or aesthetic pleasure is not the same as a desire to feel more sentiments of love or aesthetic pleasure. It would be very odd for Jane to destroy a person she loves in order to create a situation in which she would feel love for two other people whom she currently does not love; it would be odd in the same way for her to destroy one object she gets aesthetic pleasure from in order to spur others to create two such objects. The explanation for this is that the pleasurable sentiments of both love and aesthetic pleasure do not chiefly motivate "in prospect"—that is, it is not the future prospect of feeling love and aesthetic pleasure that is their main motivational influence. These feelings are, rather, directed toward the object of the affection.<sup>22</sup> That is not to say that the expectation of feeling a pleasurable sentiment can never be a motive to action, nor even to say that the thought that a course of action will cause one to feel love or aesthetic pleasure will never strengthen one's motive to take that course. The point here is merely that there is a difference between a prospective pleasurable sentiment that motivates chiefly by *pulling* the agent toward her own future pleasure, and an occurrent pleasurable sentiment that motivates chiefly by *pushing* the agent to act for the benefit of the object of the sentiment, and Hume's discussions of the motivational profiles of pride and love show that he was well aware of that difference.

Given that on Hume's pluralist account of pleasure, the indirect passions of pride, humility, love, and hatred do not imply agent-neutral maximizing trade-offs, it would hardly be surprising if his account of the moral sentiments resists that implication as well. And when we turn to Hume's texts, we do in fact find that his moral sentiments are akin to the indirect passions and aesthetic appreciation in that they are pleasures and pains that are not equivalent to, nor necessarily give rise to, the agent-neutral, maximizing desires that lead to the implausible trade-off judgments that the non-consequentialist objection foists on Hume.<sup>23</sup>

## 5. Approval of Self and Approval of Others

The differences between the motivational influences of Humean pride and love also characterize differences between the motivational influences of self-approval and approval of others. Hume himself foregrounds this connection, maintaining that “virtue and the power of producing love or pride” are “equivalent,” as are “vice and the power of producing humility or hatred” (T 3.3.1.3; SBN 575; see also 3.3.1.31; SBN 377).

On Hume’s account, approval is like pride and love in that it is not itself an immediate motive to action but rather influences the will indirectly.<sup>24</sup> There are, however, different ways approval exerts its influence. Approval of oneself influences the will in the way pride does. Approval of others influences the will in the way love does.

Hume explains how approval of oneself can influence the will in the following passage, in which he is summing up three different ways in which the virtues of benevolence benefit their possessor (I’ve added the numbers).

[1] [T]he immediate feeling of benevolence and friendship, humanity and kindness, is sweet, smooth, tender, and agreeable, independent of all fortune and accidents. [2] These virtues are besides attended with a pleasing consciousness or remembrance, and keep us in humour with ourselves as well as others; while we retain the agreeable reflection of having done our part towards mankind and society. [3] And though all men show a jealousy of our success in the pursuit of avarice and ambition; yet are we almost sure of their good-will and good-wishes, so long as we persevere in the paths of virtue, and employ ourselves in the execution of generous plans and purposes. What other passion is there where we shall find so many advantages united; [1] an agreeable sentiment, [2] a pleasing consciousness, [3] a good reputation? (E 9.2.21; SBN 282)

Hume tells us here that there are three different but converging reasons that a person with the benevolent virtues is better off: [1] the affections that are virtuous (e.g., “humanity and kindness”) are themselves pleasant; [2] one will feel the pleasure of approval toward oneself (that is., one will have “a pleasing consciousness” and a good “humour with [oneself]”) when one acts on the affections that are virtuous; and [3] one will have a “good reputation” and receive the “good-will and good-wishes” of others if one acts on the affections that are approved of as virtuous. [1] and [3] are not constituted by a person’s having feelings of approval or disapproval. It’s [2] that concerns the way approval and disapproval of one’s own character can give one a reason to be virtuous. [2] is also what Hume has in mind when he says that “in all ingenuous natures, the antipathy to treachery and

roguery is too strong to be counterbalanced by any views of profit or pecuniary advantage. Inward peace of mind, consciousness of integrity, a satisfactory review of our own conduct; these are circumstances very requisite to happiness" (E 9.2.23; SBN 283). In addition, Hume relies on the prospect view in the conclusion to the *Treatise*, when he says that acting virtuously gives us the pleasure of "inward satisfaction" and acting viciously makes us feel the pain of not being able to "bear [our own] survey" (T 3.3.6; SBN 620). That is not to say that self-approval is the only reason a person has for virtuous conduct. Indeed, the primary reason a virtuous person will act virtuously is that she has a direct or immediate motive to do what benefits others. Self-approval can nonetheless strengthen that primary motive.

The motivating influence of self-approval is, thus, like the motivating influence of pride in that it is prospective and depends on the existence of an independent, desirable motive. Suppose I am considering a future course of conduct, and I think, prospectively, that if I acted in that way I would feel approval toward myself. Hume holds that I would only feel approval because the course of conduct has some desirable feature, for example, its agreeability or usefulness to others.<sup>25</sup> So, if I think that my performance of an action would cause me to feel approval of myself, I must have a motive to perform the action that is prior to and independent of the prospect of my feeling approval toward myself. I could, for instance, want to perform an action just because it will benefit others, without having any thought whatsoever about its causing me to feel approval.<sup>26</sup> Nonetheless, it is also the case that the prospect of feeling approval can increase or give additional force to that prior desire, just as a feeling of pride can increase or give additional force to the desire to do something that will make me feel proud. Just as I will have an independent desire to possess the suit of clothes that will cause pride (since my pride is based on the clothes' having some antecedently desirable quality), so too I will have an independent desire to perform actions that I will approve of (since my approval is based on some prior desirable feature of the action). And just as the prospect of feeling pride at possessing the clothes can strengthen my motivation to possess it, so too can the prospect of approving of an action strengthen my motivation to perform it.<sup>27</sup>

Here is how the influence of self-approval can work in practice. I am considering whether to do something that will be useful to other people. I have a motive to do it because I have a desire to make other people happy. But I also have motives that pull in the opposite direction, maybe motives based on narrowly selfish desires for material gain. Failing to perform the action will, however, lead me to feel self-disapproval, and knowing this strengthens my motivation to perform the action, and thus helps to override my narrowly selfish motives to do otherwise. It is just this kind of strengthening of the motive to virtue that Hume points to when explaining why most of us will be better off eschewing sensible knavery. People with "ingenuous natures" care about the welfare of others. Their concern

for others is for the ingenuous a motive to benefit others, and their motive to benefit others will be strengthened because benefitting others will lead them to avoid self-disapproval, enabling them to experience “a satisfactory review of [their] own conduct” and a “consciousness of [their own] integrity” (E 9.2.23; SBN 283).

Now consider the influence on the will of the moral sentiments of approval and disapproval about the conduct of other people. Hume contends that my approval of a person will cause me to feel love for that person.<sup>28</sup> We have seen that Hume claims that my love for a person will also cause me to feel benevolence toward that person, or to desire her happiness. Humean psychology thus implies that when I approve of a person, I will also have a desire for that person’s happiness. And Hume makes just this point at T 3.3.1.3. and 3.3.1.31. When you come to approve (or disapprove) of someone, it will influence your will by giving rise in you to a motive to make that person happy (or miserable).

Hume refers to the motivational influence of other-approval in his discussion of “dexterity in business.” He writes:

Here is a man, who is not remarkably defective in his social qualities; but what principally recommends him is his dexterity in business, by which he has extricated himself from the greatest difficulties, and conducted the most delicate affairs with a singular address and prudence. I find an esteem for him immediately to arise in me; His company is a satisfaction to me; and before I have any farther acquaintance with him, I wou’d rather do him a service than another, whose character is in every other respect equal, but is deficient in that particular. (T 3.3.1.25)

The crucial bit here is the claim that I would rather do a service to a person with this character trait than to a person without it. According to Hume, this is evidence that the trait is a virtue, which is to say that judging someone virtuous influences the will by causing a desire to benefit her. Similarly, when explaining the similarities between natural abilities and qualities that are more traditionally thought of as virtues, Hume says that even if we choose to “refuse to natural abilities the title of virtues, we must allow, that they procure the love and esteem of mankind; that they give a new luster to the other virtues; and that a man possess’d of them is much more intitled to our good-will and services, than one entirely void of them” (T 3.3.4.2). If I approve of someone, Hume tells us here, I will have more “good-will” toward him than I will toward someone I don’t approve of and be more inclined to do him a “service.” The motivational influence of my approval of other people is to give rise to a desire for their happiness.

We can now see that the asymmetry we’ve already noted between the influence of pride and love also exists between approval of self and approval of others. Pride and self-approval influence the will *prospectively*: we do things that will cause

pride and self-approval in the future. But love and other-approval mainly influence the will *occurrently*: we do things for other people because we currently feel love or approval toward them. While pride and self-approval strengthen one's motivation to do something that will make oneself happy, love and other-approval give rise to a new motive to benefit someone else.

With the pieces we have in place from the previous four sections—the examples from aesthetics, the non-fungibility of Humean pleasures, and the motivational asymmetry between pride and self-approval, on the one hand, and love and approval of others, on the other—we can now describe clearly the failure of the non-consequentialist objection.

## **6. The Humean Answer to the Non-Consequentialist Objection**

On Hume's view, the prospect of self-approval can strengthen one's desire to act morally, and there is nothing in Hume's view that speaks against that desire's playing a non-consequentialist role in one's moral thinking. One's wish to feel self-approval is what Hume calls an "ultimate end." It gives "immediate satisfaction" (E App 1. 20; SBN 293–94). It is not something one desires because it produces anything else, such as fungible units of pleasure *per se*. One desires it for its own sake.

Consider again Jane. She has now promised her dying father that she will donate his money to a certain cause. Jane thinks that if she breaks her promise, she will feel bad about herself, will feel shame and self-disapproval. This feature of a possible course of action—that it will make her feel shame or self-disapproval—gives her a basic or ultimate reason not to do it. This reason is not reducible to the desire for any quantifiable measure of fungible units of pleasure. It functions in just the way we would expect from a non-consequentialist consideration. Jane's wish to avoid shame and self-disapproval issues a normative call that is as "pure and simple" as anything a non-consequentialist could hope for. The fact that that wish is based on sentiment does not preclude its having that kind of normative call.<sup>29</sup>

Jane will also experience a painful sentiment—of disapproval—if she learns that someone else has broken a deathbed promise to his parent, or if she learns that two people have each broken deathbed promises to their parents. Hume does not think, however, that the unpleasantness of disapproval of others influences the will chiefly by motivating the person who would feel disapproval to act in ways that will make her feel approval instead. The main motivating influence of one's disapproval toward others is to give rise to the desire for their misery (or punishment), just as the main motivating influence of love or approval is to give rise to the desire for their happiness. On Hume's view, the *prospect* of feeling love or approval towards others is not motivationally salient, which is in stark contrast to the intense motivational salience of the prospect of self-directed feelings of pride

or approval, shame or disapproval. Therefore, it is simply not true that Hume is committed to holding that Jane will prefer to trade off one of her own unjust acts for two just acts of two other people. Explicit and central features of Hume's account of moral motivation block that implication.

It would be weird if Jane preferred to do something that made her feel humiliation so that she could cause herself to feel love for three people toward whom she is currently indifferent. It would be weird if Jane was an artist and yet always preferred to spur other people to create rather than create anything herself. And it would be weird in the same way if Jane preferred to do something that made her feel disapproval of herself so that she could cause herself to feel approval toward three other people. The desires to avoid humiliation, to create art, and to avoid self-disapproval have an agent-relative character to which Hume's account of the indirect passions and the moral sentiments is acutely sensitive. These desires are not simply thrown into some Benthamite hopper of uniform impressions of pleasure to be maximized.

Moreover, even if we allow that Jane does have some motivating desire to prevent a state of affairs in which she would feel hatred or disapproval toward two other people (that is, even if we bracket the separate point, discussed in the previous paragraph, of the different motivational influences of self-approval and other-approval), there would still be no grounds for attributing to Hume the view that the strength of that desire must be greater than the strength of her desire to avoid feeling shame and disapproval of herself. The pleasure of feeling pride or self-approval is not interchangeable with the pleasure of feeling love or approval of others. Feeling the former is not, on Hume's view, feeling a certain sum of pleasures of which feeling the latter is feeling twice as much. Feeling good about oneself is pleasurable, and feeling good about other people is also pleasurable, but they are different sentiments, not greater and lesser amounts of the same sentiment. They're apples and oranges. Comparing them is no more a purely quantitative matter (where two approvals of others necessarily have twice as much deliberative weight as one approval of self) than is comparing the enjoyment of one beautiful musical composition to the enjoyment of two good bottles of wine.

And even if we were to allow that Humean self-approval and other-approval *are* fungible (which I have been arguing against), it still would not be plausible to claim that Hume is committed to a trade-off of one act of one's own immorality for two acts of others' immorality. This is because the pain of shame and disapproval of one's own act is not one, isolated sentiment at all. Feeling shame or self-disapproval is not a discrete experience temporally clipped to a one-and-done episode that caused it. It is not a single emotional moment that pops like a firework and then disappears. It persists and spreads throughout one's mental life, as Hume was well-aware. Feeling pride or self-approval gives us a "pleasing consciousness or remembrance, and keep[s] us in humour with ourselves" (E 9.2.21; SBN 282). The

“inward peace of mind” (E 9.2.21; SBN 282) of realizing you have acted well, as well as the self-recrimination of knowledge of your own knavery, is a chronic condition.

There are thus no grounds for foisting on Hume the view that a person will invariably prefer to perform an action that makes her feel shame and self-disapproval in order to prevent two others from performing actions that will make her feel hatred and other-disapproval. We may suppose that Jane prefers that the world contain fewer immoral acts rather than more. But Jane also prefers that she herself perform fewer immoral acts rather than more. And Hume is no more committed to holding that the latter preference collapses into the former than he is to holding that the satisfaction of completing a novel is one unit of pleasure of which the pleasure of reading two novels by two other people is two units, or that feeling humiliation is one negative unit of which feeling love for two people is two positive units. Indeed, Hume’s description of the long-lasting and pervasive mental effects of the results of reviews of our own conduct explains very well how the ultimate end of feeling pride and self-approval in what we ourselves have done can override the prospect of many other more momentary pleasures.

Some may worry, however, that Hume is now susceptible to another objection, namely, that his view gives us the wrong kind of reason to treat others morally.<sup>30</sup> Don’t we think we should treat people as morality demands because that’s what’s owed them and not because it produces in us the pleasure of self-approval? If Hume can explain only the latter sort of reason and not the former, doesn’t his theory fail on that account?

Hume’s view of moral motivation is not susceptible to this wrong kind of reason charge because his view does not imply that the pleasures of self-approval and pains of self-disapproval are the only or primary reasons we have for virtuous conduct. The primary reason Jane has for doing what is moral is that it will be agreeable or beneficial or just. On Hume’s account, if Jane is virtuous, she will typically be motivated immediately by her concern for others’ welfare, the social order, and the like. Indeed, Jane will approve of her own conduct just because it is typically motivated by those other concerns. What the prospect of self-approval and self-disapproval does is strengthen her motivation to act on those other concerns, which is of course very important in cases in which different considerations exert motivational force in the opposite direction.

This also explains how the Humean view can respond to the following, related objection which Julia Annas has raised: if a person acts morally because she thinks doing so will enable her to avoid the painful experience of self-disapproval, then her motive for acting in that way will be selfish; but if her motive for an action is selfish, then she won’t be able to approve of herself for performing that action. There are several reasons this objection does not discomfit the Humean. First, the agent who acts so as to avoid self-disapproval does not thereby *fail* to act on any other, approval-inducing motive. Suppose the agent is motivated to keep her

promise or assist another because she cares about keeping her promises or about others' welfare. Suppose also that she realizes that if she fails to keep her promise or doesn't lend assistance, she will feel disapproval of herself, and this thought strengthens her motivation to perform the action. That doesn't imply that she is no longer acting on the motive to keep her promise or assist another. The first motive is still there and worthy of approval. The aversion to self-disapproval doesn't make those other motives disappear or cease functioning. Secondly, Hume does not think that actions motivated by the desire to make oneself happy are never worthy of approval. Indeed, traits that are useful or beneficial to the agent herself constitute half the catalogue of Hume's virtues, and the motive to perform actions that will prevent one from feeling disapproval of oneself quite plausibly fits into that half of his catalogue. Thirdly, an agent's being more strongly motivated to do something because she will feel self-disapproval if she does not do it need not be taken to be a conscious mental event. Her concern to keep her promises or for others' welfare may be all that she is consciously aware of. But she may still be the sort of person who will disapprove of herself if she does not act in those ways, and her being that sort of person (rather than being a knave who does not feel self-disapproval) can account for some of the strength of her motivation to act in those ways, even if she does not consciously represent to herself the prospect of feeling a painful sentiment if she fails so act. Hume himself discusses the case of someone who lacks a first-order motive to do what is virtuous but who does what is virtuous anyway because he does not want to feel the painful feeling of self-disapproval (T 3.2.1.7 SBN 479). Hume does not suggest, however, what would surely be implausible, that if such a person realized that he was acting in order to avoid self-disapproval he would thereby feel self-disapproval for his own selfishness.

Jane is considering whether to perform one action or another. She is motivated to do the first because it will make someone happy or it will prevent harm or it is just. But the second will make her a great deal of money, and that creates a countervailing motive. If Jane does the first she will approve of herself, and if she does the second she will disapprove, and this strengthens her motivation to do the first. What is Jane's reason for doing the first instead of the second? The answer might have different parts. Jane may think that the first will make someone happy or prevent harm or be just. But she may also think, "I would feel terrible about myself if I did the second." The thought that she'd feel terrible about herself if she did the second is distinct from the thought of making others happy or preventing harm or being just (even if the former depends on the latter). The thought that she would feel terrible about herself is sentiment-based and self-oriented—it is a thought about what will cause her to have painful or pleasurable experiences. Those features do not, however, make it the wrong kind of reason, at least not if our task is to account for the full array of considerations that play a role in our moral thinking. People like Jane can be motivated in part by how they

think they'll feel about themselves if they act one way rather than another. The prospects of experiencing sentiments of self-approval or self-disapproval can and do function as ultimate ends in our moral thinking. Hume's account explains why this is so.

## NOTES

- 1 John Balguy, "The Foundation of Moral Goodness," in *British Moralists 1650–1800*, volume 1, ed. D. D. Raphael (Indianapolis: Hackett Publishing Co., 1991), 402.
- 2 Balguy, "The Foundation of Moral Goodness," 403. See also 407–08.
- 3 Immanuel Kant, *Groundwork for the Metaphysics of Morals*, trans. Arnulf Zweig (Oxford and New York: Oxford University Press, 2002), 432 and 228/427. (The first number is the page of this edition; the second number is the page of the 1785 edition.)
- 4 Kant, *Groundwork*, 238/437.
- 5 Kant, *Groundwork*, 226/435.
- 6 John Rawls, *Lectures on the History of Moral Philosophy* (Cambridge, MA: Harvard University Press, 2000), 46–47.
- 7 *Ibid.*, 48–50.
- 8 Stephen Darwall, *The Second-Person Standpoint* (Cambridge, MA: Harvard University Press, 2006).
- 9 Darwall, *The Second-Person Standpoint*, 178–80.
- 10 A more fundamental objection is that the Humean sentimentalism cannot account for reasons at all—that sentiments can never fund normative (as opposed to merely explanatory) reasons for action at all. I will not address that more fundamental concern here.
- 11 Darwall, *The Second-Person Standpoint*, 191. Darwall thinks that Hume has particular problems explaining our moral obligation to justice, due to his commitment to justice's artificiality and to our approving not of acts but of character traits (188); I say a little more about this in a footnote below, but I think Darwall's criticism that Hume cannot accommodate certain kinds of non-trade-off judgments, if successful, would target Hume's moral psychology in general. That the criticism would extend that widely seems to me to be implied by the following: "on Hume's official view, as we have seen, voluntary action seeks to bring about some good. This is a source of the contemporary 'Humean' theory of motivation, according to which action always results from desires (which have some apparently good [desirable] state of the world as object). As we noted in the last chapter, however, normative acceptance is irreducibly attitude-of-a-subject-regarding rather than state-of-the-world-regarding" (191).
- 12 Darwall, *The Second-Person Standpoint*, 157.

13 Stephen Darwall, *The British Moralists and the Internal Ought* (Cambridge: Cambridge University Press, 1995), 294.

14 Another version of the non-consequentialist objection is that Hume cannot accommodate our thinking that it is sometimes right to do what is just even though more people would be made happy by doing what is unjust. (See Darwall, *British Moralists and the Internal Ought*, 315. See also Christine Korsgaard, *The Sources of Normativity* [Cambridge: Cambridge University Press, 1996], 86–89.) The best Humean response to this specifically justice-based non-consequentialist objection of which I am aware is Don Garrett's in "The First Motive to Justice: Hume's Circle Argument Squared," *Hume Studies* 33 (2007): 257–88. I have addressed the justice-based non-consequentialist objection to Hume and explained why I think it fails in Michael B. Gill, "A Philosopher in his Closet," *Canadian Journal of Philosophy* 22 (1996): 231–56 and *The British Moralists on Human Nature and the Birth of Secular Ethics* (Cambridge: Cambridge University Press, 2006), 321–23. I am assuming in this paper that the justice-based objection can be met—that Hume can explain our approving of a just act even in a case in which more happiness would be created by doing the unjust thing. The different version of the non-consequentialist objection that I will be focusing on here is that Hume cannot explain why a person would prefer to perform one act of which she approves herself rather than perform an act of which she disapproves but that spurs other people to perform multiple acts of which she approves.

15 For further discussion of how consequentialists can capture all the seemingly deontological moral phenomena through the use of agent-relative desires, see James Dreier, "Structures of Normative Theories," *The Monist* 76 (1993): 22–40 and Douglas Portmore, "Consequentializing," *Philosophy Compass* 4 (2009): 329–47. For evidence that commonsense morality opts for consequentialist maximizing in the agent-relative or intrapersonal case but not in the agent-neutral or interpersonal case, see Theresa Lopez et al., "Side Constraints and the Structure of Commonsense Ethics," *Philosophical Perspectives* 32 (2009): 305–19.

16 Drawing this connection between aesthetics and morals is especially apt in the case of eighteenth-century sentimentalists such as Hume. For connecting beauty and morality was the leading idea of the entire classical sentimentalist project, with Shaftesbury, Hutcheson, Hume, and Smith all placing great emphasis on the similarity, if not the identity, of their accounts of beauty and virtue. See Michael B. Gill, "Moral Rationalism vs. Moral Sentimentalism," *Philosophy Compass* 2 (2007): 16–30.

17 Bentham's most famous expression of this view is his claim that "the game of pushpin is of equal value with the arts and sciences of music and poetry" in *The Rationale of Reward* (London: Robert Heward, 1830), 206.

18 References to the *Treatise* are to *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2007), abbreviated "T" in the text and cited by Book, part, section, and paragraph number, followed by the page number from *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd ed., revised by P. H. Nidditch (Oxford: Clarendon Press, 1978), abbreviated "SBN" in the text.

19 References to the *Enquiry* are to *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Oxford University Press, 1998), abbreviated "E" in the text and cited by section and paragraph, followed by the page number in *Enquiries Concerning*

*the Principles of Human Understanding and Concerning the Principles of Morals*, ed. L. A. Selby-Bigge, 3rd ed., revised by P. H. Nidditch (Oxford University Press, 1975), abbreviated “SBN” in the text. The appendices from the *Enquiry* are referred to as “App,” with the numerals following denoting appendix number and paragraph.

20 See Michael B. Gill, “Humean Moral Pluralism,” *History of Philosophy Quarterly* 28 (2011): 45–64.

21 None of Hume’s indirect passions of pride, humility, love, and hatred have a world-to-mind direction of fit. They are not wants, or desires that the world be a certain way. They are evaluations of the way the world is. The same is true of Hume’s moral passions of approval and disapproval. Each of these passions, however, either gives rise to or increases certain kinds of desires or wants. For fuller discussion of Hume’s belief that the indirect passions are not motives to action, see Pall Ardal, *Passions and Value in Hume’s Treatise* (Edinburgh: Edinburgh University Press, 1966), 126; Annette Baier, “The Ambiguous Limits of Desire,” in *The Ways of Desire*, ed. Joel Marks (Chicago: Chicago University Press, 1986), 53–56; Jane McIntyre, “Hume’s Passions: Direct and Indirect,” *Hume Studies* 26 (2000): 77–86; and Rachel Cohon, “Hume’s Indirect Passions,” in *A Companion to Hume*, ed. Elizabeth S. Radcliffe (Wiley-Blackwell, 2011), 159–84.

22 See Cohon, “Hume’s Indirect Passions,” 171.

23 Are the moral sentiments indirect passions? I don’t want to enter into that debate here. All that’s important for my purposes is that the moral sentiments are the same as the indirect passions with regard to their relation to motive. For discussion of this issue, see Elizabeth Radcliffe, “How Does the Humean Sense of Duty Motivate?” *Journal of the History of Philosophy* 34.3: 383–407, 1996 and Cohon, “Hume’s Indirect Passions,” 159–84.

24 See Radcliffe, “How Does the Humean Sense of Duty Motivate?” and Cohon, “Hume’s Indirect Passions.”

25 It is actually more complicated than what I say here: I will approve of the action because I think it is of a kind that is usually generated by a motive that generally produces actions that are pleasurable or useful to the agent or others.

26 Hume also discusses the unusual case of a person who lacks the primary motive to virtue but acts in the way the virtuous would because he realizes he approves of such action (T 3.2.1). For discussion of this case, see Radcliffe, “How Does the Humean Sense of Duty Motivate?,” 394–403.

27 Charlotte Brown, in “Is Hume an Internalist?” *Journal of the History of Philosophy* 26 (1988): 69–87, argues that Hume’s motivational argument against the moral rationalists (T 3.1.1.5–11; SBN 456–59) requires that moral sentiments motivate completely on their own. On this basis she maintains that the kind of view of moral motivation I present here is insufficient to fund the first premise of Hume’s motivational argument against the moral rationalists. I disagree with this aspect of Brown’s interpretation, although I will not discuss that point here. But even if Brown is right in thinking that Hume’s 3.1.1 argument requires that the moral sentiments influence the will in some kind of non-prospective way, that does not affect my claim here that Hume’s account also includes a prospective kind of moral motivation that accounts for our non-trade off judgments (and blocks the non-consequentialist objection).

28 Where love for Hume can include the relatively weak and impersonal feeling of esteem (T 2.1.5.1, 2.2.2.10, T 3.3.4.14; SBN 285, 337, 613–14).

29 This does not mean the desire for self-approval cannot be compared with other considerations, or that it will necessarily outweigh them. But it is not a quantitative, commensurable comparison. That something is a non-consequentialist consideration and that it does not reduce to a common measure with some other consideration does not mean that it can't be compared to or outweighed by another consideration. For rich discussions of the comparison of incommensurables, see Ruth Chang, *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1998).

30 Darwall makes this “wrong kind of reason” objection in *Second-Person Standpoint*, 94 and 192.