



## **Hume's Progressive View of Human Nature**

Michael Gill

*Hume Studies* Volume XXVI, Number 1 (April, 2000) 87-108.

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](mailto:humestudies-info@humesociety.org)

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

# Hume's Progressive View of Human Nature

MICHAEL GILL

In the introduction to the *Treatise*, Hume maintains that scientific advance will come only through an accurate and comprehensive conception of human nature. He praises “some late philosophers in *England*, who have begun to put the science of man on a new footing” and declares his intention to build upon their work (T xvii).<sup>1</sup>

How much of the “science of man” that Hume goes on to develop is a recapitulation of the work of the other British philosophers and how much is new? When is Hume borrowing the insights of those who came before and when is he innovating? It is difficult to answer these questions, and not just because the rules of attribution in the eighteenth century were looser than in ours. For at times the verve with which Hume writes can lead one to think that he is in the grip of a new discovery, when he is in fact recounting the ideas of a predecessor. And at other times Hume puts others' ideas to work in a manner that they themselves never considered or would have actively opposed.

There can be no doubt, however, that Hume does put forth new ideas, and some of them, I think, must be counted real advances on what came before. In this paper I will elucidate one such advance—the development of what I will call a progressive view of human nature. This view will stand out clearly when we place Hume's *Treatise* account of the virtue of justice against the backdrop of a dispute on the origin of human sociability between Shaftesbury, Mandeville, and Hutcheson, three of the five “late philosophers in *England*.” For while a number of the pieces of Hume's account appear in the work of his

three predecessors, Hume's combination of them is novel, and in the end constitutes a significant "improvement in the science of man" (T xvii).

In the first section, I outline the dispute between Mandeville, on the one hand, and Shaftesbury and Hutcheson, on the other. In the second section, I show that there are significant respects in which Hume's account of justice is in agreement with Mandeville and in disagreement with Shaftesbury and Hutcheson. In the third section, I show that there are other significant respects in which Hume's account is in agreement with Shaftesbury and Hutcheson and in disagreement with Mandeville. In the fourth section, I explain how Hume's combination of these two different aspects results in a new and improved view of human nature, one that is dynamic or progressive where that of his predecessors was static or originalist. I conclude, in the fifth section, by noting some problems with Hume's account of justice and suggesting how attention to Hume's progressive view might mitigate them.

## I

Human beings are sociable. They seek out company, live together "in Multitudes" (Mandeville I 41),<sup>2</sup> undertake large cooperative endeavors, and act in ways that benefit others. They are not solitary creatures engaged in perpetual warfare. On this point Shaftesbury, Hutcheson, and Mandeville all agree. Where Mandeville disagrees with Shaftesbury and Hutcheson is on the origin of human sociability. For while Shaftesbury and Hutcheson believe that human sociability originates in natural benevolence, Mandeville believes that it originates in self-interest.

Shaftesbury points to our friendships and our morals as compelling evidence of our natural benevolence. We all prefer "Company" to "Solitude," he says, and "almost all our Pleasures" are built upon "mutual Converse" and "Society" (Shaftesbury's *Inquiry* 63).<sup>3</sup> The aspect of friendship that gives us the most pleasure of all, moreover, is being benevolent or "doing good" to our friends (*Characteristics* II 36).<sup>4</sup> But to be virtuous is to be benevolent or to "do good" to the human species as a whole, which is just to be a "friend of mankind" (*Characteristics* II 37). This is why being virtuous provides us with such pleasure—because it, like friendship, is the expression of our natural tendency to benefit others. Indeed, this benevolent tendency is so deeply engrained in us that we even take pleasure simply in witnessing a "generous action" (*Characteristics* II 36; see also Shaftesbury's *Inquiry* 17–18, 25–26). Given that this benevolence is so universal and trenchant, Shaftesbury maintains, we can only conclude that it is "innate" or instinctive to every human being, originating not in "art, culture, or discipline" but in "mere Nature" (*Characteristics* II 135).

Hutcheson offers a similarly benevolent view of human nature, although his presentation is more systematic than Shaftesbury's. Hutcheson argues, in

particular, that every human possesses both a public and a moral sense. The public sense is a "Determination to be pleased with the *Happiness* of others, and to be uneasy at their *Misery*" (Hutcheson's *Essay* 5),<sup>5</sup> and the moral sense is that which approves of virtue, which tends to the "Publick Advantage," and disapproves of vice, which tends to the Publick "Detriment" (Hutcheson's *Essay* 8). Hutcheson takes great pains to establish that these two senses are entirely distinct from self-interest—that self-interest could never give rise to all of our motives to benefit others nor lead us to approve of all the benevolent actions of which we do in fact approve. Hutcheson also takes great pains to establish that these two benevolent senses are "*fix'd, and real and natural*" to all humans (Hutcheson's *Inquiry* 74),<sup>6</sup> "instincts" that do not result from custom, education or habit but which have been "*implanted in our Nature*" (Hutcheson's *Inquiry* 195).

Mandeville paints a very different picture, one in which humans are motivated not by "Love to others" but almost entirely by "selfishness" (Mandeville II 178). As he puts it in his attack on "Mr. *Hutcheson*," "it is not the Care of others, but the Care of itself, which Nature has trusted and charged every individual Creature with" (Mandeville II 346). It is this self-regard, according to Mandeville, not any putative "Love of our Species," that is the cause of the "Sociableness of Man" (Mandeville II 182; see also I 4, 325, 344, 346, 364).

In support of his claim that sociability originates in selfishness, Mandeville tells a long story about how humans moved, over a period of thousands of years, from their initial savage state to the complex societies we find ourselves in today. There are three stages to this development. At the first stage, early humans band together into small groups to protect themselves from the predation of "wild Beasts" (Mandeville II 230). At the second stage, small groups form into larger groups to protect themselves from the aggressive advances of other humans, maintaining a mutual defence against "the Danger Men are in from one another" (Mandeville II 266). At neither of these two stages does love for others play any role. Self-preservation alone does all the work.

The large groups formed at the second stage are unstable, however. This is because the members of the groups are liable to attack and steal from each other, as well as to betray the agreements of mutual protection that brought them together in the first place. What are needed, then, are "Antidotes, to prevent the ill Consequences" of these selfish tendencies that are "inseparable from our Nature; which yet in themselves, without Management or Restraint, are obstructive and pernicious to Society" (Mandeville II 283). The development of a written language fills this role. For once we have "Letters," we can write down our "Laws." Written laws are the means by which we can hold people to their agreements, which is the condition for the creation of groups

that are large and tolerably stable. "Therefore the third and last Step to Society is the Invention of Letters" (Mandeville II 269).

Mandeville then describes how society, once established, grows and prospers. He argues that what powers society—what generates sociable interaction and improves everyone's standard of living—is the development of commerce and standards of politeness, honor, and shame. But commerce and the standards of politeness, honor and shame originate not in benevolence but in conventions built by and upon nothing other than self-interest (e.g., Mandeville I 42 ff. and II 341 ff.).

So while Mandeville acknowledges that people do perform acts that are sociable and benefit others, he denies that humans possess any natural sentiments of sociability and benevolence. Our tendency toward sociable and benevolent action is, rather, the by-product of an artifice others invented for their own self-interests and we promote for ours. As Mandeville puts it in a passage that could sum up his attack on Shaftesbury and Hutcheson, "What you call Natural, is evidently Artificial, and belongs to Education" (Mandeville II 270). Or as he puts it elsewhere, "My Business is to demonstrate to you, that the good Qualities Men compliment our Nature and the whole Species with, are the Result of Art and Education" (Mandeville II 306).

## II

With whom does Hume side in this dispute when he is writing the *Treatise*? Certainly, Hume is no egoist, and so at least in one important respect he is much closer to Shaftesbury and Hutcheson than he is to Mandeville. Hume also believes that humans possess "natural" virtues, many of which are inherently sociable, and this too seems to place him with Shaftesbury and Hutcheson (T 578 ff.).<sup>7</sup>

At the beginning of his account of justice, however, Hume seems to side squarely with Mandeville and against Shaftesbury and Hutcheson. For the title of the first section of Hume's discussion is "*Justice, whether a natural or artificial virtue?*" which can be taken to be an indication of his intention to enter into the dispute between Mandeville, Shaftesbury, and Hutcheson (T 477).<sup>8</sup> And Hume says flat-out that his view is that justice is "not natural" but the result of "artifice," which looks to be an unambiguously Mandevillean answer (T 477).

The "short, and, I hope, convincing argument" Hume presents for the artificiality of justice seems to confirm his Mandevilleanism (T 477). Hume begins the argument by claiming that to be virtuous is to have a certain kind of motive (T 477). Hume next maintains that a virtue is natural only if the motive of which it consists is one that humans possessed in their pre-civilized or "rude and more *natural* condition, if you are pleas'd to call such a condition natural" (T 479). If, in contrast, the virtue consists of a characteristic possessed only by

"civiliz'd" individuals as a result of their having been "train'd up according to a certain discipline and education," then it is an artificial virtue (T 479). Hume then canvasses all the possible motives people could have for performing actions we think of as just, and contends that none of these is both equivalent to the virtue of justice and present in pre-civilized humans.

The first possible motive that Hume examines is a "regard to justice, and abhorrence of villainy and knavery" (T 479). Hume acknowledges—in a passage that will be very important to our later discussion—that people do possess this "regard to justice." But he contends that only "civiliz'd" humans possess it, and that humans in their "rude and more *natural* condition" did not (T 479). Hume's argument here rests on the idea that a virtue consists of the possession of a certain kind of motive (T 477). But a "regard to justice" is simply the motive to perform the actions that would be performed by someone who possesses the virtue of justice. The person who possesses the virtue of justice, therefore, must *first* possess some motive other than the "regard to justice" before anyone else can *later* develop the (derivative) "regard to justice."<sup>9</sup> Another way of putting this point is by saying that a simple regard for justice on its own has no content; it is a *de dicto* motive, or the motive to-perform-just-actions. A simple regard for justice can arise, therefore, only *after* a person knows what the content of justice is. Thus the simple regard for justice cannot be the origin of the contentful idea of justice, since the contentful idea must pre-date the simple regard.<sup>10</sup>

The second motive for justice that Hume examines is "self-love" or "a concern for our private interest or reputation" (T 480). This motive, unlike the simple regard for justice, has existed in all humans at all times, civilized or not. But self-love is obviously not the motive that constitutes the virtue of justice. For just conduct is not always in one's self-interest, so someone motivated exclusively by self-love will at least sometimes be guilty of injustice. To the extent that self-love remains in its pre-civilized or "*uncultivated*" condition, the divergence between it and justice will be very great indeed (T 488).

The third candidate for the motive of justice that Hume considers is "the regard to publick interest" (T 480), or "public benevolence" (T 482), a desire to benefit humanity as a whole. Hume's view of this candidate is particularly important for our purposes, since Mandeville disagrees with Shaftesbury and Hutcheson precisely on the question of whether or not humans possess the motive of benevolence towards humankind. Thus Hume's rejection of this motive seems to leave little room for doubt as to which side he takes in the dispute among his three predecessors.

Hume offers three "considerations" for denying that the virtue of justice consists of a "regard to the publick interest" (T 480). The first reason is that someone motivated solely by the desire to benefit humanity as a whole would at least sometimes commit injustice, since a "single act of justice" may actually harm the "public interest," as for instance when "a man of merit . . . restores

a great fortune to a miser, or a seditious bigot" (T 497). Hume acknowledges that most acts of justice do benefit the public and that even the "single act" of restoring a great fortune to a miser or bigot can have the beneficial effect of serving as an example to those who would otherwise be tempted to commit harmful injustice. But this connection between the public interest and the example of the single act (which is harmful considered in isolation but beneficial when considered as an example) is not natural, since it will hold only within civilized societies, in which the rules of justice are already well known and established.

Hume's second "consideration," furthermore, consists of a counterexample that shows that even within a civil society with established rules of justice, the connection between the public interest and single acts of justice will not always hold. He points out, specifically, that the conditions of a loan to a miser or seditious bigot might include that the repayment of it be made in secret, "as when the lender wou'd conceal his riches," in which case the repayment could not achieve even the beneficial goal of serving as an example to others (T 481). But although in such a case there is not even an artificial convention that creates a connection between justice and the public interest, "there is no moralist, who will affirm, that the duty and the obligation ceases" (T 481). There is, then, no chance that the duty or obligation is the same thing as the motive to benefit the public.

Hume's third "consideration" for rejecting "public benevolence" as the origin of justice cuts even deeper than the first two (T 482). He begins by pointing out that it is simply an undeniable fact that people do not have "the public interest" in mind when they "pay their creditors, perform their promises, and abstain from theft, and robbery" (T 481). "That is a motive," he says, "too remote and too sublime to affect the generality of mankind, and operate with any force in actions so contrary to private interest as are frequently those of justice and common honesty" (T 481).

Hume goes on, moreover, to argue that the public interest could *never* be the motive behind acts of "justice and common honesty" because the motive of public interest—upon which Shaftesbury and Hutcheson build their entire theory of human nature—does not exist in any human, "rude" or "civiliz'd." As he puts it, "In general, it may be affirm'd, that there is no such passion in human minds, as the love of mankind, merely as such, independent of personal qualities, or services, or of relation to ourself" (T 481).

Hume's denial of the existence of a love for mankind merely as such is grounded in the psychology he develops in Book II of the *Treatise*, where he argues that love and hate are caused by the qualities individuals possess, not by their simple humanity itself (T 330). I feel love for someone, Hume maintains there, because she is associated in my mind with a quality that causes me pleasure, just as I feel hatred for someone because he is associated in my mind with a quality that causes me pain. Both people are humans, of

course, but that fact merely enables the association of impressions and ideas to work in my mind. Their humanity is merely an associative conductor of impressions and ideas, not a quality that itself can arouse either love or hatred. As Hume explains,

[m]an in general, or human nature, is nothing but the object both of love and hatred, and requires some other cause, which by a double relation of impressions and ideas, may excite these passions. . . . There are no phaenomena that point out any such kind affection to men, independent of their merit, and every other circumstance. (T 481–482)

Hume acknowledges that humans prefer being with others to being alone. But like Mandeville, he denies the Shaftesburean claim that this preference for company reveals a benevolent concern for humanity. Just as Mandeville says that most people engage in friendly leisure activity simply for their own enjoyment (Mandeville I 336–344, II 183), so too Hume maintains that we “love company in general; but 'tis as we love any other amusement” (T 482).

Hume also acknowledges that we have the capacity to feel friendship or love toward any other individual under certain circumstances. But he denies that this capacity constitutes evidence that we have the quite different capacity to feel love or friendship toward all humans under any circumstances. He writes:

An *Englishman* in *Italy* is a friend: A *European* in *China*; and perhaps a man wou'd be belov'd as such, were we to meet him in the moon. But this proceeds only from the relation to ourselves; which in these cases gathers force by being confined to a few persons. (T 482)

Hume's examples in this passage are particularly revealing, for they are clear echoes of the following passage from *The Fable of the Bees* in which Mandeville is explicitly attacking Shaftesbury's benevolent view of human nature.

Two *Londoners*, whose Business oblige them not to have any Commerce together, may know, see, and pass by one another every Day upon the *Exchange*, with not much greater Civility than Bulls would: Let them meet at *Bristol* they'll pull off their Hats, and on the least Opportunity enter into Conversation, and be glad of one another's Company. When *French*, *English* and *Dutch* meet in *China* or any other Pagan Country, being all *Europeans*, they look upon one another as Country-men, and if no Passion interferes they will feel a natural Propensity to love one another. . . . These things by superficial Judges are attributed to Man's Sociableness, his natural Propensity to

Friendship and love of Company; but whoever will duly examine things and look into Man more narrowly, will find that on all these Occasions we only endeavor to strengthen our Interest, and are moved by the Causes already alleg'd. (Mandeville I 343)

Our capacity to feel friendship toward other particular individuals, Mandeville maintains here, has misled "superficial Judges" such as Shaftesbury into attributing to humans a public benevolence, while in fact this phenomenon is better explained by other "Causes." But this is the very same point we have just seen Hume advance in his denial of the existence of "public benevolence." We find, then, that when criticizing Shaftesbury's and Hutcheson's publicly benevolent conception of human nature, Hume and Mandeville speak in one voice.

The final candidate for the original motive of justice that Hume considers is the motive of "*private benevolence, or a regard to the interests of the party concern'd*" (T 482). But Hume quickly dismisses this candidate because, first of all, we might hate the person toward whom we have an obligation of justice, and indeed the person may actually deserve the hatred of all humankind, but our obligation to him does not diminish, nonetheless. Hume also points out that the person toward whom we have an obligation of justice may be a "miser" who "can make no use of what I wou'd deprive him of," or a "profligate debauchee" who "wou'd rather receive harm than benefit" from my giving him what he is owed (T 482). Once again, however, our obligation to give the person what he is owed would remain the same, even though our giving it to him does not benefit him at all. So our desire to benefit a particular person, or "private benevolence," cannot be the origin of the virtue of justice.

To sum up, then: According to Hume, if our sense of justice is natural, it will have at its foundation a natural sentiment. There is, however, no such natural sentiment to be found. All the possible sentimental candidates are either nonexistent (public benevolence), non-natural (a regard to the virtue of justice), or incompatible with some instances of our obligation to justice (self-love and private benevolence [as well as public benevolence, if it did exist]). Hume concludes, therefore, that "the sense of justice and injustice is not deriv'd from nature, but arises artificially" (T 483)—a result that is conspicuously Mandevillean.

Hume's constructive account of the development of society and the artifice of justice is also basically Mandevillean in that it, too, is grounded almost entirely in self-interest.<sup>11</sup> Hume begins his account by noting the relative helplessness of individual humans to meet their needs and satisfy their desires.<sup>12</sup> As he explains,

Of all the animals, with which this globe is peopled, there is none towards whom nature seems, at first sight, to have exercis'd more

cruelty than towards man, in the numberless wants and necessities, with which she has loaded him, and in the slender means, which she affords to the relieving these necessities. . . . In man alone, this unnatural conjunction of infirmity, and of necessity, may be observ'd in its greatest perfection. (T 484–485)

Mandeville is impressed by the same point, maintaining that humans are uniquely “curs'd” by “Obstacles” and that “[a]ll the Element are our Enemies” (Mandeville I 344–345; see also Mandeville I 205).

Hume and Mandeville both then go on to argue that humans develop societies just because it is the only way for them to overcome their relative helplessness. As Mandeville puts it, “The Love Man has for his Ease and Security, and his perpetual Desire of meliorating his Condition, must be sufficient Motives to make him fond of Society; considering the necessitous and helpless Condition of his Nature” (Mandeville II 180; see also Mandeville I 344). And as Hume has it,

'Tis by society alone he is able to supply his defects. . . . By society all his infirmities are compensated; and tho' in that situation his wants multiply every moment upon him, yet his abilities are still more augmented, and leave him in every respect more satisfied and happy, than 'tis possible for him, in his savage and solitary condition, ever to become. (T 485)

Here Hume attributes “the origin of society” to “self-interest,” for, as he explains, “[t]here is no passion . . . capable of controlling the interested affection, but the very affection itself, by an alteration of its direction” (T 492). We must take the antecedent of this view to be Mandeville, not Shaftesbury or Hutcheson.

Clearly Mandevillian as well is Hume's story of how justice originates in conventions grounded in “*selfishness and limited generosity*” (T 494). Everyone wants to secure his or her “goods,” Hume tells us. But certain of those goods—namely, “possessions . . . we have acquir'd by our industry and good fortune”—are inherently insecure, since other people both can “ravish” them and have motive to do so (T 487). Humans come to realize, however, that they will all stand to benefit if they enter into a stabilizing “artifice” or “convention” whereby each person leaves every other person “in the peaceable enjoyment of what he may acquire by his fortune and industry” (T 489). Thus

I observe, that it will be for my interest to leave another in the possession of his goods, *provided* he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually express'd,

and is known to both, it produces a suitable resolution and behaviour.  
(T 490)

Hume's account of this mutual agreement to refrain from taking each other's goods is fascinating and compelling. Hume argues, in particular, that this agreement is one that does not rely on "the interposition of a promise" but "arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it" (T 490). It is as a result of this gradual agreement that there "arise the ideas of justice and injustice" (T 490). This progressive agreement, in other words, is the very origin of justice.

Now, what is most crucial about this story for our purposes is that in it Hume very clearly attributes the artifice that is the origin of justice to "the love of gain" and not to natural benevolence (T 492). The artifice, Hume says, is grounded entirely in a "sense of interest" (T 490). Indeed, a few pages after his story of the development of the artifice of justice, Hume explicitly claims that justice would never have come into existence if, instead of being selfish, "every man had a tender regard for another" (T 494). "Encrease to a sufficient degree the benevolence of man," Hume maintains, "and you render justice useless" (T 494–495). For the very purpose of justice is to "restrain" the "selfishness of man," and it is thus the selfishness of man in which justice originates (T 495).

I should not overemphasize the similarity between Mandeville and Hume on the origins of society and justice, for Hume clearly thinks that even in their "rude" and "uncultivated" state people are not as exclusively self-interested as Mandeville does (T 486–487).<sup>13</sup> But Hume does insist that societies and the conventions of justice develop as a result of the redirection of self-interest, and this is an unmistakably Mandevillean position. As Hume sums it up, it is "*only from the selfishness and confin'd generosity of men, along with the scanty provision nature has made for his wants, that justice derives its origin*" (T 495).

### III

Within Hume's account of justice, however, there is also one very conspicuous point at which he appears to side with Shaftesbury and Hutcheson and against Mandeville. This point concerns the sincerity of persons' commitment to the virtue of justice.

Mandeville insists that the vast majority of people who pretend to be virtuous are in fact hypocrites (e.g., Mandeville I 331). Indeed, one of the most characteristic features of Mandeville's writings as a whole is the view that most of what passes in society for virtue is really just a counterfeit (e.g., Mandeville I 254)—that although we all claim to occupy a moral high ground, our overriding motives are actually almost entirely selfish and thus possess no moral worth. As Mandeville puts it, "There is not a quarter of the Wisdom, solid

Knowledge, or intrinsick Worth, in the World, that Men talk of, and compliment one another with; and of Virtue or Religion there is not an hundredth Part in Reality of what there is in Appearance" (Mandeville II 340).

One of the most characteristic features of the writings of both Shaftesbury and Hutcheson, in contrast, is a much more positive view of human nature, one that allows that many people at least some of the time do really and truly care about what is virtuous. Central to Shaftesbury and Hutcheson's arguments for this more positive view of human nature is the contention that people such as Hobbes and Mandeville fail miserably in their attempts to find selfishness at the heart of all human behavior.

Does Hume think people really and truly care about the virtue of justice and are not simply hypocritical moral counterfeiters? He clearly does. Does he think the virtue of justice is distinct from self-interest? He clearly thinks that as well.

We have seen that Hume believes that "in his rude and more *natural* condition," a human would reject "as perfectly unintelligible" the idea of repaying a loan simply out of a "regard to justice" (T 479). But Hume also believes that this idea is entirely intelligible to people within society, and that in fact a "regard to justice" does play a role in the conduct of "civiliz'd" people. As he puts it,

I ask, *What reason or motive have I to restore the money?* It will, perhaps, be said, that my regard to justice, and abhorrence of villainy and knavery, are sufficient reasons for me, if I have the least grain of honesty, or sense of duty and obligation. And this answer, no doubt, is just and satisfactory to man in his civiliz'd state, and when train'd up according to a certain discipline and education. (T 479)

We find, then, that at this point in the *Treatise*, in the midst of his Mandevillian argument for the artificiality of justice, Hume attributes to civilized people the "antipathy to treachery and roguery" that he will later rely upon in the *Enquiry* in his much more Hutchesonian account of the reason to be just (EPM 283).

Hume's reference to a "sense of duty and obligation" is not, moreover, an isolated comment. In the next section of the *Treatise*, when addressing himself to the question of why "*we annex the idea of virtue to justice, and of vice to injustice,*" he explains in detail how we acquire this sense of duty, or "regard to justice" (T 498). People initially care about justice only because it accords with self-interest, he tells us there. But over time, they develop mental associations that lead them to approve of justice even when it does not promote their self-interest, and to disapprove of injustice even when it does promote their self-interest.<sup>14</sup> These unselfish approvals and disapprovals constitute not a counterfeit concern for the virtue of justice, but the real thing. Hume argues,

in other words, that as a result of an associative “progress of sentiments” (T 500), people who originally had only selfish reasons for caring about justice eventually come to possess a commitment to justice that can run “contrary to private interest” (T 481).<sup>15</sup>

If there is any lingering doubt about where Hume stands in the debate between Shaftesbury, Hutcheson, and Mandeville on the sincerity of persons’ commitment to the virtue of justice, his discussion of the “artifice of politicians” should dispel it. In that discussion, Hume acknowledges that the commitment to justice has been “forwarded by the artifice of politicians, who, in order to govern men more easily, and preserve peace in human society, have endeavour’d to produce an esteem for justice, and an abhorrence of injustice” (T 500). Hume then maintains, however, that this point “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” (T 500). The implication here, of course, is that there *is* a sense of virtue among mankind. This is significant for our purposes because in making this claim Hume is clearly signaling his disagreement with Mandeville and his agreement with Shaftesbury and Hutcheson. We can be sure this is Hume’s aim, since Mandeville was famous for claiming that “the Distinction between *Virtue* and *Vice*” was “the Contrivance of Politicians” (Mandeville I 50–51), and Hume says in this passage that the “writers on morals” are led astray by their false belief that the “artifice of politicians” is “the sole cause of the distinction we make betwixt vice and virtue” (T 500; see also T 533 and T 578–580).<sup>16</sup> When attacking Mandeville on this point, moreover, Hume is obviously seconding Hutcheson, who scorned Mandeville’s contention that people act virtuously only because they have been manipulated by the “*Statues and Panegyrics*” of “cunning Governours” (Hutcheson’s *Inquiry* II 121).

#### IV

So Hume claims, first, that justice is an artificial virtue that originates in self-interest. And he also claims, second, that people really do exhibit the non-self-interested virtue of justice. Mandeville would agree with the first claim, Shaftesbury and Hutcheson would agree with the second, but none of them would accept the possibility that both could be true. Indeed, all of the British moralists who preceded Hume would have thought that these two claims are incompatible. Hume’s two-part position on justice is thus something significantly new in British moral philosophy.

What is significantly new in Hume’s account will stand out clearly when set against a point on which Mandeville, Shaftesbury, and Hutcheson all agree. Shaftesbury, Hutcheson, and Mandeville all believe [1] that if Mandeville is right about the self-interested origins of human sociability, then most of what passes for virtue is really just hypocrisy or counterfeit. This is because all three

of them believe [2] that if Mandeville is right about the self-interested origins of sociability, then there really is no difference between the motives of the saint and the motives of the knave. They believe, that is, that if human sociability originates in self-interest, then the seeming benevolence of the saint will really be just as selfish as the mischief of the knave. Now [1] does follow plausibly enough from [2], for if everyone is equally selfish, then the person who pretends to care about others simply for their own sake will be trading in counterfeit virtue. If no distinction can be drawn between the motives of the saint and the motives of the knave, then there will be a very important sense in which virtue is not real.<sup>17</sup>

But why think [2] is true? Why think that the claim that humans were originally motivated to become sociable because it served their self-interest implies the claim that self-interest still remains humans' only motive for sociable behavior?

I think that Shaftesbury, Hutcheson, and Mandeville all believed [2] because of their implicit adherence to what we may call a static or *originalist* view of human nature. According to this view, humans' original motives always remain their only truly fundamental ones. The basic elements of human motivation are fixed. Experience and socialization can alter the focus or direction of the original human motives, but they cannot create a new kind of motive altogether.<sup>18</sup> The ultimate driving forces of human conduct stay the same. Thus, on the static or originalist view, if humans were in the past initially motivated to become sociable because it served their self-interest, then it must be the case that their motivation to continue to be sociable in the present is also self-interested. And if humans are in the present motivated to be sociable for non-selfish reasons, then it must be the case that they were in the past motivated to become sociable because of some non-selfish reasons. What is impossible, according to this view, is that an original selfish motive to become sociable could be supplemented and even contravened by a non-selfish motive that did not exist before sociability emerged.

This originalist view of human nature is implicit not only in Shaftesbury, Mandeville, and Hutcheson but also in thinkers otherwise as diverse as Hobbes and Cudworth. Hobbes, for instance, maintains that in the state of nature humans' chief motivation is self-interest—that self-interest is far and away the most predominant original human motive. And it is this self-interest that leads humans to form a commonwealth. Now the situation in the commonwealth is very different from the state of nature, and so the focus or direction of self-interest will be different as well. There are, for instance, rules that are in one's interest to follow in the commonwealth that would not be in one's interest to follow in the state of nature. Hobbes even suggests at times that once we are in the commonwealth we ought to reconceive our entire way of thinking about how to promote our self-interest. But self-interest always remains the underlying motivation for Hobbes. For Hobbes, that is, the fundamental

human drive in the commonwealth is the same as the fundamental human drive in the state of nature.<sup>19</sup>

Cudworth, of course, adamantly opposes Hobbes's position, holding that humans possess non-selfish motives to perform actions that are "fit" or in line with the immutable and eternal principles of morality. What is important for our purposes, however, is that Cudworth takes great pains to establish that these moral motives are innate to human nature—that they have not developed over time, but have always been fully present in the mind of every human that has ever lived. Indeed, Cudworth believes that we can be moral agents only if our moral motives are innate. And he believes this because he thinks that truly moral motives could never be generated from non-moral ingredients. If we did not originally possess truly moral motives, according to Cudworth, we never would. If a certain kind of motive is not innate to human nature, then it will not be able to exist in human nature at all.<sup>20</sup>

Hume's predecessors believed, then, that we were stuck with our original motives, and that we could change human behavior only by changing the circumstances in which those original motives operated. They thought that if people possessed a sincere and non-selfish regard for justice at the present time, then humans must always have possessed a sincere and non-selfish regard for justice; and that if people in the past did not possess a sincere and non-selfish regard for justice, then people could not possess such a regard at the present time.

In contrast to this originalist view of human nature, Hume holds what we can call a dynamic or *progressive* view, one which allows that original motives can evolve into other motives of different kinds. Hume believes we can develop new motives, ones that were not part of our original endowment. And what he reveals in his account of the virtue of justice is how—as a result of "a progress of sentiments"—a new motive can develop.<sup>21</sup> He shows, that is, how there can evolve a regard for justice that is not original but is nonetheless entirely sincere—how a real commitment to the impartiality of justice grows out of our originally partial nature.<sup>22</sup>

The key to Hume's account of the evolution of a regard to justice is his use of the principles of association; it is the principles of association that move us from an uncultivated partiality that is "confin'd to ourselves" (T 489) to a cultivated impartiality that encompasses the "*public interest*" (T 500). Hume explains this movement by first pointing out that unjust acts generally cause more harm than good. Our awareness of this feature of most unjust acts, coupled with our sympathetically grounded disposition (itself a kind of association) to disapprove of that which harms others, leads us to disapprove of unjust acts that do not affect our own interest. But if we have represented to us enough harmful acts of injustice that do not affect our own interests, and if (as we must) we feel disapproval in most of these cases, we will eventually develop the associative habit of conjoining disapproval and injustice. Once

this associative habit develops, we will tend to feel disapproval toward all unjust acts, even those that benefit us. This is because of our addiction to the "general rule," or associative tendency to overgeneralize, which causes our disapproval of the injustice of others to become connected in our minds to the injustice we commit ourselves (T 499).<sup>23</sup>

Hume uses the principles of association, moreover, to explain not just the regard to justice but also the evolution of almost all of the other morally interesting human phenomena (such as pride, love, approval, and the general points of view). The principles of association, in other words, are the very mechanisms that power the Humean progress of sentiments. Indeed, it is his progressive use of the principles of association that might be Hume's greatest innovation. For while Locke and Hutcheson had earlier noted the human mind's tendency to associate, they had held that the resulting associations were merely unfortunate mental kinks, anomalous perversions of what was true and right of the human constitution. But in Hume's view of human nature, association takes center-stage. Associative progressions, according to Hume, give rise to just those things that Locke and Hutcheson tried to locate in the mind's original pre-associative state.

Attention to the word "original" and its cognates may help bring what is importantly new about Hume's position into sharper focus. In the decades prior to the publication of the *Treatise*, many moralists addressed themselves to the question of the "origin of morals," and their answers often involved claims about what was and what was not an "original" principle of human nature.<sup>24</sup> But what did they mean by "original"? We can separate out two different senses or ways of understanding originality: a chronological one and a nonchronological one.

On the chronological understanding, something is original if nothing else existed before it. So when we ask about the "origin of morals" in the chronological sense, we are asking about the genealogical history or earliest causes of morality. This question is analogous to the question of, say, the origin of the game of chess, an answer to which would involve a discussion of other earlier games with different rules and how they evolved into the particular game of chess we know today.

On the nonchronological understanding, in contrast, something is original if it is the source of or reason for something else. So when we ask about the "origin" of morals in the nonchronological sense, we are asking about the source of, or underlying reason for, our moral judgments. This question is analogous to the question of, say, the origin of the authority of our elected officials, an answer to which might cite the power of the people in a democratic system of government.

In the dispute between Mandeville, Shaftesbury, and Hutcheson, these chronological and nonchronological senses are conflated. This is because the

three of them implicitly assume that the chronological origins of human sociability will also be the source of, or underlying reason for, our current sociability. Indeed, virtually all of the British moralists who preceded Hume conflated the chronological and nonchronological "origins" of morality. Virtually all of them assumed that the chronologically earliest cause of human morality would also be the source of, or the underlying reason for, our current moral judgments.<sup>25</sup>

Hume, however, is plainly aware of the difference between chronological and nonchronological accounts. Nowhere is this more apparent than in his theory of justice, in which he explicitly distinguishes the historical causes of the institution of justice from the reasons people currently have for their "regard to justice." As he puts it, "*Thus self-interest is the original motive to the establishment of justice; but a sympathy with public interest is the source of the moral approbation which attends that virtue*" (T 499–500; see also T 533 and T 577–580). To arrive at this statement would have been impossible for Hume's predecessors, because from their originalist perspective they could not see the difference between the motive to establish justice and the source of our moral commitment to justice. But from Hume's progressive view of human nature, the difference is clear.<sup>26</sup>

## V

Hume's account of justice is not without its problems, however. For parts of it appear to conflict both with claims Hume makes elsewhere and with other parts of Hume's account of justice itself. Two internal problems in particular stand out.

First of all, while Hume implies in some passages that our moral approval of justice can directly motivate us to just action even when it is not in our interest (T 479), he maintains in other passages that our moral approval of justice "is too weak to controul our Passions" (T 670). Indeed, at times Hume seems to endorse the general position that moral approval, of which our "regard to justice" is an example, cannot motivate at all.<sup>27</sup> Now, even if Hume in the end holds that the virtue of justice never controls our actions, his position will still be distinct from Mandeville's, since Hume does clearly hold that our moral judgments of just and unjust actions—including our own—are truly nonselfish. But the position that nonselfishness influences only our evaluations and not our conduct does shade closer to Mandeville than do a number of Hume's *Treatise* comments about the "regard to justice, and abhorrence of villainy and knavery" (T 479) and his *Enquiry* response to the sensible knave.

The second problem is that Hume's account of our moral approval of justice seems to conflict with his claim that "the ultimate object of our praise and approbation" is "motive" (T 477). This is because it seems that Hume himself has already established that there is no motive that is equivalent to the virtue

of justice; that certainly seems to be the implication, anyway, of the argument we examined above in which Hume rules out the motives of selfishness, public benevolence and private benevolence. We also saw, of course, that Hume believes that civilized people possess a "regard to justice," but (unless Hume has established a "sophistry" by "reasoning in a circle" [T 483]) *that* cannot be the motive that is the "ultimate object of our praise and approbation," since *that* is itself a kind of approbation that must have something distinct from it as its object.<sup>28</sup> (The first problem could reappear here as well, since to hold that the "regard to justice" is the "motive" of which we approve would be to imply that an approval can also be a motive.)

Darwall, whose exploration of these problems is the deepest of which I am aware, maintains that "no interpretation can dissolve" them entirely.<sup>29</sup> I am inclined to agree with him, if what he means is that it is impossible to combine all of Hume's statements on justice into one perfectly coherent whole. At the same time, I think that attention to Hume's progressive view and its relationship to his predecessors mitigates the problems somewhat. For both of the problems can be traced back to an originalist idea—or a piece of excess originalist baggage—that makes occasional appearances in Hume's text but is not organic to his deepest philosophical insights.

The offending originalist idea is that the shape and function of every human "passion" and "sense" is sharply defined and fixed by our original constitution, unsusceptible to radical alteration, innovation, or supplementation. This idea is particularly conspicuous in Hutcheson, who erects a firewall between, on the one hand, motivating passions that are approved of ("*Motives or Desires*" that lead to "*Election*") and, on the other hand, moral "*Approbation*," which does not motivate (Hutcheson's *Essay* 208–209).

It is entirely reasonable, however, to think that Hume's deepest philosophical insights are incompatible with this originalist idea, as is evident not only in his progressive account of justice but also by his profoundly anti-Hutchesonian associative account of the passions, according to which "one passion will always be mixt and confounded with the other" (T 441).<sup>30</sup> But once this originalist idea is jettisoned, the way will be cleared both for an account of how approvals that originally have no motivational force can evolve into motivations (which would dissolve the first problem), and for an account of how selfish endorsements of a large subset of a certain type of action can evolve into moral approvals of that type of action in general (which would dissolve the second problem).<sup>31</sup> Originalist thinking will of course resist the possibility of such instances of sentimental evolution.<sup>32</sup> But what I have been trying to show here is that it is just this originalist thinking beyond which Hume is trying to move.

Some of the things Hume says about justice do appear to contradict other of the things he says. But there are principled reasons for taking certain of Hume's claims to be central to his thought and for taking others to be more or

less dispensable intellectual inheritances. The conflicts between the former and the latter are almost inevitable—cracks that result from pouring new wine into old skins.

Philosophical ideas are never *sui generis*, cut entirely out of whole cloth. But there are new philosophical ideas nonetheless, and some of them are real advances on what came before. My aim here has been to elucidate one such advance—to show how Hume fashioned from materials found in the works of his originalist predecessors a new progressive view of human nature.

## NOTES

A shorter version of this paper was presented to the 27th International Hume Society Conference in Williamsburg, Virginia, July 2000. Many thanks to Rachel Cohon, who commented on that paper and whose "Hume's Difficulty with the Virtue of Honesty" (*Hume Studies* 23.1 [1997]: 91–112) does a superb job of making many of the points I try to make here.

1. David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd ed. revised by P. H. Nidditch (Oxford: Clarendon Press, 1978), hereafter referred to as "T."
2. Bernard Mandeville, *The Fable of the Bees*, ed. F. B. Kaye, 2 vols. (originally published Oxford: Clarendon Press, 1924; reissued Indianapolis: LibertyClassics, 1988). All references to this work will be cited in the text by "Mandeville," with the first (roman) numeral denoting the volume number and the second (arabic) numeral denoting the page number.
3. Anthony Ashley Cooper, Third Earl of Shaftesbury, *An Inquiry concerning Virtue, or Merit* (Manchester: Manchester University Press, 1977). All references to this work will be noted in the body of the text by "Shaftesbury's *Inquiry*."
4. Anthony Ashley Cooper, Third Earl of Shaftesbury, *Characteristics of Men, Manners, Opinions, Times* (Indianapolis: Bobbs-Merrill, 1964). All references to this work will be noted in the body of the text by "*Characteristics*," with the first (roman) numeral denoting the volume and the second (arabic) numeral denoting the page.
5. Francis Hutcheson, *An Essay on the Nature and Conduct of the Passions and Affections with Illustrations on the Moral Sense*, 3rd ed. (originally published in London, 1742; reissued Gainesville, Fla.: Scholars' Facsimiles & Reprints, 1969). All references to this work will be noted by "Hutcheson's *Essay*."
6. Francis Hutcheson, *An Inquiry into the Original of our Ideas of Beauty and Virtue* (originally published London, 1725; reissued Hildesheim: Georg Olms, 1971). All references to this work will be noted by "Hutcheson's *Inquiry*."
7. Although even Hume's account of the natural virtues, insofar as they rely on the associative principles of the mind, is more of a departure from Shaftesbury and Hutcheson than it might initially appear.

8. Hutcheson's most extensive discussions of justice occur in works published after Hume's *Treatise*, so we should not take Hume's *Treatise* account of justice to be engaged specifically with Hutcheson's account of justice. But we may take Hume's *Treatise* account of justice to be engaged with Hutcheson's general view of morals and human nature, as expounded in the early works Hume had read.

9. But see what I call the "second problem" in section V below for complications of this argument.

10. Perhaps Hume moves too quickly here, in that he doesn't consider the possibility that all humans—civilized or not—are capable of determining the content of justice through the use of reason alone, and that therefore a simple regard for justice and an idea of justice's content can arise simultaneously (or even in a sense be the same thing). Mackie is very possibly right as well when he contends that in this discussion Hume simply denies without much argument the possibility that there was an original instinct to respect property, keep promises, and in general follow the rules of justice (John Mackie, *Hume's Moral Theory* [London: Routledge & Kegan Paul, 1980], 80–81).

11. Commentators such as Mackie (*Hume's Moral Theory*, 83) and Haakonssen (Knud Haakonssen, *The Science of a Legislator* [Cambridge: Cambridge University Press, 1981], 21) have noted Mandeville's influence on Hume's constructive account of justice, although I think they tend to understate it somewhat.

12. Hume's constructive account of justice is a conjectural history. For an excellent account of the Scottish context of this kind of genetic explanation, see Roger L. Emerson, "Conjectural History and Scottish Philosophers," *Historical Papers 1984* (Ottawa: Canadian Historical Association, 1984), 63–90, as well as John B. Stewart, *Opinion and Reform in Hume's Political Philosophy* (Princeton: Princeton University Press, 1992), 152–193.

13. Baier is especially clear on this difference—on how Hume's pre-civil human beings are naturally much nicer than Mandeville's (Annette Baier, *A Progress of Sentiments* [Cambridge, Mass.: Harvard University Press, 1991], 222).

14. See also T 672, where Hume says that the "inventors of [the laws of nations, of modesty, and of good-manners] had chiefly in view their own Interest. But we carry our Approbation of them into the most distant Countreys and Ages, and much beyond our own Interest." Because Hume acknowledges that justice can be counter to self-interest, I think Gauthier is wrong to try to shoehorn moral obligation into self-interested obligation (David Gauthier, "David Hume, Contractarian," *Philosophical Review* 88 [1979]: 26–29). I also think Baier is on questionable ground in claiming that "Hume seems to require that, for something to be a moral obligation, it must first satisfy the test of self-interest which convention imposes" (*A Progress of Sentiments*, 243).

15. Emerson, in his discussion of Scottish conjectural histories, tells us, "Development and progress were used synonymously by eighteenth century Scots for whom progress did not usually imply a necessarily better state but only a change" ("Conjectural History and Scottish Philosophers," 65). It is interesting to ask whether Hume thought that small communities in which there was no (need of the virtue of) justice were in a better or worse state than larger nations in which there is (the need of the virtue of) justice. I tend to think Hume would not have been greatly interested in this question, focusing instead on how the development

occurred and not on whether it was for good or ill. In any event, Emerson's comment is an important reminder that just because Hume holds what I call a "progressive view of human nature," it does not mean that Hume was committed to the Whiggish view that human nature was necessarily advancing or improving.

16. Baron suggests, I think, that we take this passage with a grain of salt, and read Hume as in the end advancing the Mandevillean idea that our sense of virtue is created (albeit from fairly benevolent motives) by the noble lie of politicians (Marcia Baron, "Hume's Noble Lie: An Account of his Artificial Virtues," *Canadian Journal of Philosophy* 12 [1982]: 539–555).

17. See David Fate Norton, *David Hume, Common-Sense Moralist, Sceptical Metaphysician* (Princeton: Princeton University Press, 1982), 55–93, for a discussion of "moral realism" in this period.

18. One of Shaftesbury's main topics in the *Inquiry* is to determine how various religious beliefs affect a person's moral sentiments, although he doesn't seem to think that there can arise new kinds of sentiments altogether. Hutcheson discusses at great length how original human passions can become corrupted, although he does not seem to think a new salutary passion can ever be created. (In Michael B. Gill, "Nature and Association in the Moral Theory of Francis Hutcheson," *History of Philosophy Quarterly* 12.3 [July 1995]: 281–302, I discuss Hutcheson's view of how passions can become corrupted and how to prevent them from becoming so.) In his later works (particularly the Dialogues in Part II of *Fable of the Bees* and *An Enquiry into the Origin of Honour*), Mandeville gives a more subtle gradualist (Kaye calls it "evolutionary") account of civilizing developments than he might appear to endorse in his earlier works (such as "The Grumbling Hive," and "An Enquiry into the Origin of Moral Virtue"), although he never strays far from his fundamentally egoistic view. For discussion of this issue in Mandeville, see Kaye's Introduction in Mandeville I lxiv–lxvi.

19. For an excellent discussion of this issue in Hobbes, see Hampton, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1986), 27–57.

20. See *A Treatise concerning Eternal and Immutable Morality*, ed. Sarah Hutton (Cambridge: Cambridge University Press, 1996), 144–145.

21. Baier's title, *A Progress of Sentiments*, comes of course from this passage in the text, and throughout her book she explains Hume's claims about how humans develop and grow various capacities and judgments that were not part of their original uncultivated endowment. Stewart makes similar points throughout his discussion of Hume's "conjectural history," as does Streeter when he speaks of Hume's "account of socialization" (Ryan Streeter, "Hume and the Origins of Justice," *Dialogue* 38 [October 1995]: 11) and of "the possibility of a progressional moral cultivation in people" (*ibid.*, 8). I believe that Feiser also describes accurately the role of artifice in the development of the motive to act justly, although he overlooks the role artifice also plays in the development of our approval of (even non-beneficial acts) of justice (James Feiser, "Hume's Motivational Distinction between Natural and Artificial Virtues," *British Journal of the History of Philosophy* 5 [1997]: 373–388). See also Mackie, *Hume's Moral Theory*, 87.

22. I believe that Stroud misses somewhat the progressive aspect of Hume's view of human nature, which leads him to conclude that Hume fails to account adequately for our sense of justice (Barry Stroud, *Hume* [London: Routledge, 1977], 204–210).

It seems to me that Stroud notes accurately the generally Mandevillean starting point of Hume's account but doesn't allow that Hume has the progressivist resources to explain how we can move beyond these selfish beginnings.

23. This paragraph is largely repeated from p. 37 of Michael B. Gill, "Fantastic Associations and Addictive General Rules: A Fundamental Difference between Hutcheson and Hume," *Hume Studies* 22.1 (1996): 23–48. See that paper for a fuller discussion of Hume's innovative use of the principles of association to explain numerous moral phenomena.

24. Mandeville, for instance, writes "*An Enquiry into the Origin of Moral Virtue*" and Hutcheson writes "*An Inquiry into the Original of our Ideas of Beauty and Virtue.*"

25. Similar ambiguities infect the words "ground," "foundation," "derive," and "principle."

26. No one appreciated Hume's anti-originalist progressive view of human nature more fully than Mill, who develops the same idea in chapter 4 of *Utilitarianism* when he explains how virtue, money, fame, and the like can come to be desired as ends even though they were not "originally and naturally" so desired. He says, for instance, that "the strongest natural attraction . . . of power and of fame, is the immense aid they give to the attainment of our other wishes; and it is the strong association thus generated between them and all our objects of desire, which gives to the direct desire of them the intensity it often assumes, so as in some characters to surpass in strength all other desires" (John Stuart Mill, *Utilitarianism*, ed. George Sher [Indianapolis: Hackett, 1979], 36). He also says of virtue that there "was no original desire of it, or motive to it, save its conduciveness to pleasure, and especially to protection from pain. But through the association thus formed, it may be felt a good in itself, and desired as such with as great intensity as any other good" (*Utilitarianism*, 37). See also Mill's discussion of conscience (*Utilitarianism*, 27–28) and moral feelings (*Utilitarianism*, 30–32). Another philosopher with a clearly progressive view of human nature is Rousseau.

27. See Brown's "Is Hume an Internalist?" (Charlotte Brown, *Journal of the History of Philosophy* 26 [1988]: 69–87) for an important discussion of this issue. Darwall (Stephen Darwall, *The British Moralists and the Internal "Ought"* [Cambridge: Cambridge University Press, 1995], 304–305) and Schneewind (J. B. Schneewind, *The Invention of Autonomy* [Cambridge: Cambridge University Press, 1998], 370) also highlight this problem.

28. I thus find compelling both Mackie's claim that to make Hume's "view coherent, we must, I think, take him to be relaxing the principles that actions count as virtuous *only* in so far as they are signs of virtuous motives in the case of the artificial virtues" (*Hume's Moral Theory*, 80) and Darwall's claim that we must attribute to Hume the idea of a "rule obligation," which is distinct from the natural and moral obligations explicitly mentioned in the *Treatise* (*The British Moralists*, 315).

29. Darwall, *The British Moralists*, 290.

30. For an excellent discussion of the complex nature of Hume's inheritance from Hutcheson, see especially James Moore, "Hume and Hutcheson," in M. A. Stewart and John P. Wright (eds.), *Hume and Hume's Connexions* (Edinburgh: Edinburgh University Press, 1994), 23–57. I discuss how Hume's associative account of the passions constitutes a profound difference with Hutcheson in my "Fantastic Associations and Addictive General Rules."

31. Of course, it is not enough just to clear the way for such accounts. The accounts themselves have to be plausibly developed. Gauthier, Baron and Haakonssen all argue that in order to work, Hume's accounts must include the idea that at some point we commit the "error" of believing that there is a natural motive to justice when in fact there is none. (Baron, as noted above, also attributes to Hume the view that politicians and educators must lie to the people in order to cause them to make this error.) My own view is that the key to Hume's explanation of the development of what Darwall calls our "rule obligation" to justice is Hume's concept of our addiction to general rules, which might appear to be a kind of error but is not necessarily so. See also Ainslie, who explores the idea that the problems with Hume's account might be solved through attention to Hume's idea of "national characters" (Donald C. Ainslie, "The Problem of the National Self in Hume's Theory of Justice," *Hume Studies* 21.2 [November 1995]: 289–313).

32. An originalist bias might also be part of the anti-naturalist positions recently argued for by Hampton (Jean Hampton, *The Authority of Reason* [Cambridge: Cambridge University Press, 1998]) and Korsgaard (Christine M. Korsgaard, *The Sources of Normativity* [Cambridge: Cambridge University Press, 1994]). In Michael B. Gill, "Reflexivity and Justification in Hume's Moral Theory," *Canadian Journal of Philosophy* 26.2 (1996): 231–256, I discuss Korsgaard's criticisms of Hume's naturalist account of justice.