



## **Artificial Virtues and the Sensible Knave**

David Gauthier

*Hume Studies* Volume XVIII, Number 2 (November, 1992) 401-428.

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/>

# Artificial Virtues and the Sensible Knave<sup>1</sup>

David Gauthier

## I

Hume's account in the *Treatise* of the artificial virtues, their obligation and motivation, resists easy interpretation. Two passages, taken from his discussion of promises, will introduce the problems I propose to examine. First:

No action can be requir'd of us as our duty, unless there be implanted in human nature some actuating passion or motive, capable of producing the action. This motive cannot be the sense of duty. A sense of duty supposes an antecedent obligation: And where an action is not requir'd by any natural passion, it cannot be requir'd by any natural obligation; since it may be omitted without proving any defect or imperfection in the mind and temper, and consequently without any vice. Now 'tis evident we have no motive leading us to the performance of promises, distinct from a sense of duty. If we thought, that promises had no moral obligation, we never shou'd feel any inclination to observe them. ... But as there is naturally no inclination to observe promises, distinct from a sense of their obligation; it follows, that fidelity is no natural virtue, and that promises have no force, antecedent to human conventions.<sup>2</sup>

Ignore for the present the last sentence quoted. Hume seems clearly to be claiming that (1) an action can be a duty only if there is some motive to perform it other than a sense of duty, and (2) there is no motive to perform an act of promise-keeping other than a sense of duty. From (1) and (2) we may infer that (3) an act of promise-keeping can not be a duty. But since the sense of duty motivates an action only if it is apprehended to be a duty, then from (2) and (3) we may conclude that (4) there is no motive to perform an act of promise-keeping that does not involve the *mistaken* apprehension that it is a duty.

And the second passage:

Afterwards a sentiment of morals concurs with interest, and becomes a new obligation upon mankind. ... The difficulties,

that occur to us, in supposing a moral obligation to attend promises, we either surmount or elude. For instance; the expression of a resolution is not commonly suppos'd to be obligatory; and we cannot readily conceive how the making use of a certain form of words shou'd be able to cause any material difference. Here, therefore, we *feign* a new act of the mind, which we call the *willing* an obligation; and on this we suppose the morality to depend. But we have prov'd already, that there is no such act of the mind, and consequently that promises impose no natural obligation. (T 523)

This time, ignore for the present the first sentence. The remainder would seem to confirm my reading of the first passage. Hume claims that (5) there is no natural obligation to perform an act of promise-keeping, but (6) we feign the willing of such an obligation. If, as Hume has claimed in the previous passage, (2) there is no motivation to perform an act of promise-keeping other than a sense of duty, and if obligation and duty may be taken as equivalent, then we may conclude as before that (4) there is no motive to perform an act of promise-keeping that does not involve the *mistaken* apprehension that it is a duty.

In his theory of the artificial virtues and vices, Hume considers "the three fundamental laws of nature, *that of the stability of possession, of its transference by consent, and of the performance of promises*" (T 526). It would seem, from the passages cited, that Hume must suppose that there is no real obligation to keep promises, but only a feigned obligation. And since the reasoning that Hume employs in the first of the passages that I have cited is explicitly identified with "that reasoning, which prov'd justice in general to be an artificial virtue" (T 518), it seems clear that he must also suppose that there is no real obligation to uphold stability of possession and its transference by consent—no real obligation to observe the other laws of justice or equity.<sup>3</sup> We may confirm this inference by examining a third passage, in which Hume claims to establish the artificiality of justice:

we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance; and as no action can be equitable or meritorious, where it cannot arise from some separate motive, there is here an evident sophistry and reasoning in a circle. Unless, therefore, we will allow, that nature has establish'd a sophistry, and render'd it necessary and unavoidable, we must allow, that the sense of justice and injustice is not deriv'd from

nature, but arises artificially, tho' necessarily from education, and human conventions. (T 483)

Hume claims that (7) an action can be equitable only if there is some motive to perform it other than a sense of its equity, and (8) there is no universal motive to perform an act of equity other than its equity. These correspond to (1) and (2) in his account of promise-keeping. And from them we may infer, ignoring for the present the complications introduced in the final sentence of the passage, that (9) the apprehension of an action as equitable is a *misapprehension*, and (10) the motive to perform an act as equitable involves this misapprehension. But how could a real obligation to observe the laws of equity rest on a misapprehension?

Hume does not pursue a concern with sophistry, or insist on the presence of a feigned act, in his discussion of the other artificial virtues, allegiance and chastity, but his treatment of them is sufficiently parallel with his account of equity and fidelity that we may, I think, assume that he might readily have introduced such a concern. And so it would seem that Hume's theory of the artificial virtues and vices, of justice and fidelity, and by plausible extension allegiance and chastity, is an *error theory*.<sup>4</sup> Hume all but acknowledges this feature when he says,

As the obligation of promises is an invention for the interest of society, 'tis warp'd into as many different forms as that interest requires, and even runs into direct contradictions, rather than lose sight of its object. (T 524)

The interest of society requires that we believe ourselves under an obligation to keep our promises, and so we are induced to have such a belief. This part of morality is a salutary error; Hume does not wish to undermine its hold on us. And so he does not draw out the implications of his arguments. But they are present, and apparent enough when we enquire into the foundations of the virtues that Hume treats as artificial, even if we conveniently forget them in those of our daily interactions that depend on strict adherence to the supposed duties of justice, fidelity, and allegiance—and perhaps also on adherence to chastity, although I shall henceforth ignore that supposed virtue.

I propose now to follow two lines of enquiry. The first offers a different reading of Hume, distinguishing the natural obligation to justice, fidelity or promise-keeping, and allegiance from the moral obligation, and showing that the natural obligation is not feigned, and that the moral obligation may be related to it. This reading distinguishes Hume's theory of the artificial virtues from his theory of

morality and relegates the latter to a subordinate position. The second line of enquiry begins by introducing the sensible knave from the *Enquiry Concerning the Principles of Morals*. The knave offers an objection, not to this different reading of Hume as a reading, but to the adequacy of the theory so read. I shall consider how he might be answered, if not by the Hume of the *Enquiry* then by the Hume of the *Treatise*, and this will return us to an error theory that undermines the artificial virtues. I shall conclude by considering whether Hume may have an alternative answer that would leave these virtues intact.

## II

1. Hume makes seemingly contradictory claims about both our motive to promise-keeping and our motive to equity. I have already quoted his insistence that "we have no motive leading us to the performance of promises, distinct from a sense of duty" (T 518), and "that we have naturally no real or universal motive for observing the laws of equity, but the very equity and merit of that observance" (T 483). And yet in numerous passages he insists that there is such a motive—a motive of interest.

To the imposition then, and observance of these rules [that is, of justice], both in general, and in every particular instance, they are at first mov'd only by a regard to interest; and this motive, on the first formation of society, is sufficiently strong and forcible. (T 499)

And again:

The same self-love, therefore, which renders men so incommodious to each other, taking a new and more convenient direction, produces the rules of justice, and is the *first* motive of their observance. (T 543)

And with respect to our "interest in the *institution* and *observance* of promises," Hume says that all that is needed for the convention of promising is

that every one have a sense of interest in the faithful fulfilling of engagements, and express that sense to other members of the society. This immediately causes that interest to operate upon them; and interest is the *first* obligation to the performance of promises. (T 522-23)

The second of the passages with which I began follows on immediately; thus its reference to “a new obligation.”

Not only does Hume make seemingly contradictory claims about motivation, he also makes seemingly contradictory claims about obligation. He insists “that promises impose no natural obligation,” but only after claiming that “interest is the *first* obligation to the performance of promises” (T 523). And the obligation of interest is natural, as Hume makes clear when he speaks of the “*natural* obligation to justice, *viz.* interest” (T 498), and “the *natural* obligations of interest” referring specifically both to promises and to allegiance (T 545). So there is a natural obligation to the performance of promises, just as there is a natural obligation attending each of the artificial virtues. Can we reconcile Hume’s claims into a consistent doctrine?<sup>5</sup>

Hume acknowledges that he uses ‘natural’ sometimes to contrast with ‘artificial’, sometimes with ‘moral’ (and also with ‘civil’, although that will not concern me) (T 474-75). Now we may read Hume’s claims about the obligation to keep promises as (i) a denial that promises oblige naturally *as opposed to artificially*, and (ii) an insistence that promises oblige naturally *as well as* (and indeed prior to) *morally*. Not only may Hume consistently hold, but I shall argue that he does actually hold, that the interested obligation to keep promises is in itself non-moral, and so by contrast natural, and that it is also artificial, and so by contrast not natural. Hume would have avoided some of his apparent contradictions had he not chosen to use ‘natural’ to make both of these quite different contrasts—which at this point I have only introduced and not explained.

Hume’s claims about motivation are not so readily accommodated. Here we must note that the claims may be differentiated by context. Hume begins his discussion by considering what motive we might have to the performance of particular just acts, and dismisses in turn the agent’s interest, his concern for the public interest, and private benevolence (T 480-83). And he begins his discussion of promises again by considering particular acts, and argues both that promise-making is in itself unintelligible, and that were it intelligible, promise-keeping, like abstinence from injustice, would lack any natural motive. Contextualized, Hume’s claim is then that we have no sufficient motive to the performance of the particular acts required by justice and fidelity, *if* we consider those acts in abstraction from the *practices* that they exemplify. And this is to say that we have no *natural* motive sufficient to the performance of these acts.

But the practices do provide us with a self-interested motive that is at least initially sufficient. Consider, first, justice narrowly conceived, or equity. Each person, aware of his need for society to afford him “additional *force, ability, and security*” in obtaining and enjoying

external goods (T 485) and recognizing that, "As the improvement ... of these goods is the chief advantage of society, so the *instability* of their possession ... is the chief impediment" (T 488), must seek a remedy in

a convention enter'd into by all the members of the society to bestow stability on the possession of those external goods, and leave every one in the peaceable enjoyment of what he may acquire by his fortune and industry. (T 489)

Such a convention restrains the "heedless and impetuous movement" (T 489) of the love of gain,

since 'tis evident, that the passion is much better satisfy'd by its restraint, than by its liberty, and that by preserving society, we make much greater advances in the acquiring possessions, than by running into the solitary and forlorn condition, which must follow upon violence and an universal licence. (T 492)

And so the love of gain, suitably redirected by this reflection, restrains itself:

'tis impossible for men to consult their interest in so effectual a manner, as by an universal and inflexible observance of the rules of justice, by which alone they can preserve society. (T 534)

This restraint is, we might say, an *artificial* motive, since it arises only as part of a conventional practice.

A particular act of justice may be contrary to the agent's interest, to the public interest, to benevolence—to everything that might be thought possible to motivate it as an isolated act. But

this momentary ill is amply compensated by the steady prosecution of the rule, and by the peace and order, which it establishes in society. And even every individual person must find himself a gainer, on ballancing the account. (T 497)

Each, then, has an interest in the rule of justice, which he

expresses ... to his fellows, along with the resolution he has taken of squaring his actions by it, on condition that others will do the same. ... And thus justice establishes itself by a kind of convention or agreement; that is, by a sense of interest, suppos'd to be common to all, and where every single act is

## ARTIFICIAL VIRTUES AND THE SENSIBLE KNAVE

perform'd in expectation that others are to perform the like.  
(T 498)

Hume is insistent both on the convention and on the expectation by each person that others will adhere to it "since nothing but this combination can render justice advantageous, or afford me any motives to conform my self to its rules" (T 498).

Hume argues that each person must expect to gain from the universal observance of the rules of justice, and so each has a sufficient motive for his own observance. What is the link that I have expressed by 'and so'? More specifically, how does the advantage of general observance motivate particular observance, since as Hume admits, "Taking any single act, my justice may be pernicious in every respect" (T 498)? This concession must be read as implicitly qualified; taking any single act *apart from its place in the general practice*, my justice may be pernicious. Thus Hume says, "A single act of justice ... *were it to stand alone*, ... may, in itself, be very prejudicial to society" (T 497, emphasis added). But "disorder and confusion follow upon *every* breach of these rules [of justice]" (T 499, emphasis added). In the context of the general practice I make "the supposition, that others are to imitate my example" (T 498), so that each particular failure undermines general observance. Each person must expect every choice that he makes between conforming to and violating the rules of justice to have an effect on the behaviour of others with consequences for his own advantage sufficient to afford him with a normally adequate motive for conformity. Nothing less will meet the requirements of Hume's argument. His own questioning of the plausibility of this expectation, expressed through the sensible knave, will afford the transition to the second part of my discussion, but for the present I shall let the argument go unchallenged.

2. Hume repeats the essentials of his account of our interest in justice or equity in establishing our interest in both fidelity and allegiance. The general practice must be invoked to warrant the particular acts, or indeed in the case of promising, to make them intelligible. In promising I express a resolution to perform the act promised, and I do so by invoking "a *certain form of words*," invented so that "we might give each other security of our conduct in any particular incident" (T 522). Merely expressing a resolution to act would give me no new motive or obligation, but the convention invoked by the use of the language of promising gives rise to such a motive, in subjecting me "to the penalty of never being trusted again in case of failure" (T 522). And so Hume concludes that "whoever uses [these words] is immediately bound by his interest to execute his engagements" (T 522).

Precisely what is the interest that motivates promise-keeping?<sup>6</sup> Hume seems to introduce two quite different interests, although he makes no distinction between them: the interest each of us has in the practice of promising, which makes possible "mutual trust and confidence in the common offices of life" (T 544) and the interest each of us has in himself being trusted. Each of these interests takes us beyond the context of a particular promise, but in significantly different ways. The first is parallel to one's interest in maintaining the rules of justice, which would be undermined by violating them. The second is parallel to the interest one would have in not being ostracized were one to violate those rules. In the first case one seeks to avoid a general weakening of the practice of trust, and one may suppose that the adverse consequences of such weakening would fall as heavily on one's fellows as on oneself. In the second case one seeks to avoid being excluded from the benefits of the practice, and one need not suppose that the adverse consequences of such exclusion would fall significantly on others, or that one's act of promise-breaking would have any material effect on the general practice of trust.

I have compared one's interest in being trusted with the interest one would have in not being ostracized were one to violate the rules of justice. Hume speaks as if not being trusted were a penalty, and we might think that ostracism would also be a penalty, but an interest in avoiding such a *penalty* is quite different from the interest one would have in avoiding *punishment*. A person who is not trusted, or who is ostracized, is unable to invoke the practice when it would be in her interest to do so. Ostracism denies access to the benefits of participation. Punishment imposes costs. Punishment for violating a practice is a *second* artifice, invented to reinforce the effectiveness of the original, which is the practice itself. It can motivate conformity whether or not persons value the practice to which they conform. Ostracism involves no second artifice, but only the exclusion of certain persons from the scope of the original. It can motivate conformity only if persons value the practice.

It is important to distinguish generally between a person's interest in the maintenance of a practice and her interest in her participation in the practice. Hume appeals to both kinds of interest in his discussion of fidelity, although not in his discussion of justice, but as I have said, he does not seem to be aware of the distinction. I shall want to appeal to it later after the sensible knave has appeared on the scene.

3. Let us suppose that Hume has successfully shown that each person has an interest in the practices associated with the artificial virtues sufficient normally to motivate her performance of particular acts of justice, fidelity, and (although I have not examined Hume's discussion

of it) allegiance. What is the basis of the further claim that each has a *natural obligation* to these acts, considered in relation to the practice they exemplify? Indeed, what is a natural obligation? Recall first that Hume distinguishes 'natural' from 'artificial', and also from 'moral'. A natural obligation, I have already suggested, is natural as opposed to moral, but not as opposed to artificial, since it arises in the context of artifice—of a conventional practice of justice, or promising, or allegiance. But still, what is it?

Hume does not tell us. Speaking of the "*natural* obligation to justice, *viz.* interest," he assures us that it "has been fully explain'd" (T 498), but the explanation is simply the account that I have already sketched of our interest in preserving society in order better to acquire external goods and to live with our fellows in peace and order. He tells us that "interest is the *first* obligation to the performance of promises" (T 523), and speaks later of "the *natural* obligations of interest, which are distinct in promises and allegiance" (T 545). And in each of the places that I have just quoted, Hume proceeds to speak of a different, *moral* obligation, thus confirming my claim that he is using 'natural' to contrast with 'moral'. But in each of these passages he simply takes for granted the appropriateness of speaking of an obligation, while offering no account of what distinguishes a motive that is obligatory from one that is not. And while it is possible to construct an account of *moral* obligation from what he says, as I shall endeavour to do presently, this account depends on Hume's view of morality in ways that make it unsuitable to try to apply it to explain natural obligation.

Before proceeding I should again emphasize that not all of Hume's references to natural obligation are to be interpreted by contrasting 'natural' with 'moral'. When Hume insists "that promises impose no natural obligation" (T 523), the intended contrast must be with obligation arising from human practices based on voluntary conventions, and so with *artificial* obligation. Hume denies that there are or can be natural (that is, non-artificial) obligations associated with justice, fidelity, and allegiance, but he asserts that there are natural (that is, non-moral) obligations of interest to each of these virtues. My concern here is to fill the void created by Hume's omission of any explication of these interested obligations.

I shall therefore simply propose an account of natural (non-moral) obligation that seems to me to fit Hume's use and to have some independent merit. I take my cue here from Hume's insistence that the motive to justice requires a redirection of interest.

There is no passion, therefore, capable of controlling the interested affection, but the very affection itself, by an alteration of its direction. Now this alteration must

necessarily take place upon the least reflection; since 'tis evident, that the passion is much better satisfy'd by its restraint, than by its liberty. (T 492)

My proposal is that the motive associated with a redirected passion constitutes a natural obligation. The redirection curbs the natural force of the passion, restraining it so that it may be "better satisfy'd." But we may suppose that the natural tendency of the passion to its unrestrained fulfilment is still present, so that the effect of the redirection is felt as *restraint*. The agent recognizes her interest as requiring restraint, but nevertheless what is required is restraint, and so she takes herself to be under an obligation. There is nothing moral in this obligation, since it relates solely to the better satisfaction of her own love of gain, but the absence of morality is not equivalent to the absence of restraint.

The absence of morality should, however, be emphasized before we consider how Hume introduces it. He insists that adherence to the rules of justice is "at first mov'd only by a regard to interest; and this motive, on the first formation of society, is sufficiently strong and forcible" (T 499). However, as society increases, "this interest is more remote; nor do men so readily perceive, that disorder and confusion follow upon every breach of these rules, as in a more narrow and contracted society," and so "we may frequently lose sight of that interest, which we have in maintaining order" (T 499). Note that Hume does not say that as society increases, its members cease to have an interest sufficient to motivate adherence to the rules of justice. Quite the contrary; Hume insists that in society the "consequences of every breach of equity seem to lie very remote," so that they "are not able to counterbalance any immediate advantage, that may be reap'd from it," yet they are "never the less real for being remote" (T 535). It is therefore the short-sightedness of persons, in failing to recognize the effects of their actions on maintaining the rules of justice, that makes interest motivationally insufficient in practice. Were persons as aware of the more remote and diffuse effects of their actions as of the immediate effects, there would presumably be no need for any additional motivation to reinforce their redirected interest in assuring its own restraint.

Morality is therefore practically necessary but theoretically dispensable in motivating conformity to the requirements of justice, fidelity, and allegiance. Now this is not an original observation. Bernard Wand emphasizes this feature of Hume's account as an objection to it:

We never consider that the function of moral obligation is to serve as a substitute for narrow self-interest. ... we consider

that when we act from moral motives, the nature of our motives is radically different from that of self-interest.<sup>7</sup>

Perhaps we do think this, but perhaps we are confused in thinking it. I shall, however, argue later that Hume should not treat morality as theoretically dispensable, even though he is right to see it, if not as a substitute, then perhaps as a replacement for narrow, or even broad, self-interest. But before turning to this argument—which again awaits the sensible knave—I shall examine Hume's view of the motivational role played by morality in relation to the artificial virtues, and his idea of moral obligation.

4. When I consider a character trait, I have an idea of its immediate agreeableness or disagreeableness, and of its utility or disutility, both to its possessor and to other persons. This awareness gives rise to an impression in me, corresponding to my idea of its effects; this is the work of sympathy. And this sympathetic impression is the basis for my moral assessment of the trait. Thus if I consider someone's benevolence, I think of the benefits she confers on those with whom she associates, and this idea of its utility gives rise to a pleasing impression which leads me to judge her benevolence to be morally virtuous. In like manner, if I consider someone's disposition to justice or equity, I think of the service afforded to the public interest by the universal practice to which her disposition is related, and this idea then gives rise to a pleasing impression leading me to judge the justice of her character to be morally virtuous. Fidelity, or a regard to promises, and allegiance, or obedience to civil authority, similarly receive my moral approbation.

A more careful and detailed account would reveal problems in Hume's view of moral approbation that I want simply to bypass as not directly relevant to my main concerns. I shall focus on the relation between moral approbation and moral obligation. Hume says,

when any action, or quality of the mind, pleases us *after a certain manner*, we say it is virtuous; and when the neglect, or non-performance of it, displeases us *after a like manner*, we say that we lie under an obligation to perform it. (T 517)

To understand Hume's position we must distinguish the standpoint of the agent from that of the spectator. And I shall do this by considering first a natural virtue, one of those common motivational traits that receive moral approbation independent of conventional practices—generosity. I observe the gifts you make to your friends and to charitable causes, and I take your behaviour to indicate a generous disposition. The idea of such a disposition pleases me *after a certain*

*manner*; this pleasure in turn leads me to judge that your disposition is a virtue and that you are virtuous. Or I observe your refusal to contribute to charitable causes, and the mean and paltry gifts you make to friends, and I take your behaviour to indicate niggardliness and the lack of a generous disposition. The idea of such a lack displeases me *after a like manner*, and this displeasure leads me to judge that your niggardliness is a vice, and that you lie under an obligation to give more generously.

This is the spectator's view, but it requires certain assumptions about the agent. For the last judgement, of obligation, presupposes that persons are naturally moved to generous behaviour. I have quoted Hume as insisting that

where an action is not requir'd by any natural passion, it cannot be requir'd by any natural obligation; since it may be omitted without proving any defect or imperfection in the mind and temper, and consequently without any vice. (T 518)

Here 'natural' contrasts with 'artificial', and so what Hume says must apply to the moral obligation associated with the natural virtue of generosity. If I take you to be under an obligation to give more generously, then I must suppose that such giving is required by a natural passion, and that its absence indicates a defect in your temper.

But if you lack generosity, then how can you be motivated to more generous behaviour? Hume has an answer:

When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle, may hate himself upon that account, and may perform the action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least, to disguise to himself, as much as possible, his want of it. (T 479)

This capacity, to hate oneself for lacking virtue and consequently to act to acquire or at least simulate it, is present to motivate obligatory behaviour should its natural motive be lacking. Thus the spectator, in claiming that the agent has an obligation to act generously, ascribes to her either the natural passion of generosity or the self-hatred occasioned by its absence. And it is this second ascription, of self-hatred at the absence of the natural motive, that is essential to *moral* obligation.

This account explains the moral obligation to behave in accordance with a natural virtue. Note that it requires that the virtue be identified

initially as a character trait that motivates in a manner quite independent of moral approbation, and that any agent may be supposed to be motivated either by possession of this trait or by the felt lack of it. But is a similar explanation possible for the moral obligation to behave in accordance with an artificial virtue such as justice? The sense of justice is not a natural trait, whose absence is a defect in one's temper; rather it "arises artificially, tho' necessarily from education, and human conventions" (T 483). Can it be identified as a character trait that motivates in a manner quite independent of moral approbation? A person has an interested motive to the performance of just acts, and this motive is independent of moral approbation, but does it arise from the sense of justice? If not, and if the only motive arising from the sense of justice is the moral obligation to just acts, then we can not explain this obligation by appealing to self-hatred occasioned by the felt lack of a prior virtuous motive.

I suggest, however, that the interested motive to justice does indeed arise from what we come to identify as the sense of justice. In itself the sense of justice is simply self-love, or more precisely what Hume calls "the interested affection," or "love of gain" (T 492), but *redirected* towards its fuller satisfaction through its own restraint by the conventionally instituted laws of society. Absent these conventions there could be no sense of justice, since there would be no basis for distinguishing particular actions as just or unjust, or for identifying any behaviour as part of a general practice that offers increased satisfaction through restraint.

The moral obligation to be just, then, arises from reflection on the beneficial tendencies of redirected interest. I observe the effects of the universal practice of justice, and I take your adherence to the practice to indicate a just disposition. The idea of such a disposition pleases me *after a certain manner*; this pleasure in turn leads me to judge that your disposition is a virtue, and that you are virtuous. Or I observe your disregard of justice and the disorder and confusion that your conduct occasions and I take your disregard to indicate the lack of a just disposition. The idea of such a lack displeases me *after a like manner*, and this displeasure leads me to judge that your injustice is a vice, and that you lie under an obligation to act justly. If you feel yourself devoid of the *redirected* interest that constitutes justice, you may hate yourself upon that account, and may perform the just action without the motive from a certain sense of duty, in order to acquire by practice a sense of justice, or at least, to disguise to yourself, as much as possible, your want of it. This capacity, to hate oneself for lacking justice and consequently to act to acquire or at least simulate it, is present to motivate obligatory behaviour should its interested motive be lacking.

If I am right, then the explanation of the moral obligation to behave in accordance with an artificial virtue is strictly parallel to the explanation of the moral obligation to behave in accordance with a natural virtue. It requires that the artificial virtue be identified initially as redirected interest, which of course motivates in a manner quite independent of moral approbation, and that any agent may be supposed to be motivated either by possession of this interest or by the self-hatred occasioned by the felt lack of it.

In his account of Hume's theory of justice, Knud Haakonssen notes the problem that Hume faces in specifying the motive that "is missing if we behave unjustly" and that "we come to hate ourselves for not having."<sup>8</sup> The only actual motive is, as I have argued, self-interest, which Hume does not regard as a character trait that is morally approved, but rather as morally neutral. And Haakonssen is then led to the view that Hume supposes that people

*imagine* that there is a natural motive (and thus a character trait) ... namely the willing of an obligation. And it is this *imagined* motive that, through sympathy, they come to approve of. And when they find that, for very good reasons [presumably its non-existence], they do not have this motive themselves, they come to hate themselves, and this self-hatred creates the sense of duty.<sup>9</sup>

In effect, then, Haakonssen claims that even though the only motive truly arising from the sense of justice is the moral obligation to just acts, yet we can not explain this obligation by appealing to self-hatred occasioned by the felt lack of a prior motive, because we *feign* this prior motive. This is an ingenious argument, to which the sensible knave will force us to return. But it seems to me that given the interested obligation to justice, Hume is able to offer a more satisfying account of the virtuous motive that we take to be commonly present, and whose absence gives rise to self-hatred, in terms of *redirected* self-interest. Self-interest itself is indeed, as Haakonssen maintains, morally neutral, but redirected towards the conventions on which society depends, it is transformed into a virtue.

5. This completes my attempt to read Hume's account of the obligations to equity, fidelity and allegiance. In each case there is a convention or conventions, embraced by each person to increase her acquisitions and her security. Although adherence to these conventions requires each person to restrain her interest, this restraint is mandated by the interest itself, redirected by reflection on the greater satisfaction that it attains by maintaining the conventions. Natural (non-moral)

obligation is based on this redirected interest, and this obligation is initially effective in maintaining social order. But as society becomes more complex, persons fail sufficiently to appreciate the force of their interested obligations, which require, in practice if not in theory, to be supplemented by morality (and also, of course, by magistracy). The beneficial effects of social conventions lead, through sympathy, to treating the redirected interests that uphold them as virtues, and so to moral obligations to justice, fidelity, and allegiance that add to the force of the initial, interested obligations. And so we may agree with Hume, when he ends "[s]ome further reflections concerning justice and injustice" by saying,

Upon the whole, then, we are to consider this distinction betwixt justice and injustice, as having two different foundations, *viz.* that of *self-interest*, when men observe, that 'tis impossible to live in society without restraining themselves by certain rules; and that of *morality*, when this interest is once observ'd to be common to all mankind, and men receive a pleasure from the view of such actions as tend to the peace of society, and an uneasiness from such as are contrary to it. 'Tis the voluntary convention and artifice of men, which makes the first interest take place; and therefore those laws of justice are so far to be consider'd as *artificial*. After that interest is once establish'd and acknowledg'd, the sense of morality in the observance of these rules follows *naturally*, and of itself. (T 533)

I began by quoting three passages from Hume. I read the first as claiming on the face of it that (1) an action can be a duty only if there is some motive to perform it other than a sense of duty, and (2) there is no motive to perform an act of promise-keeping other than a sense of duty, from which it would follow that (3) an act of promise-keeping can not be a duty, and (4) there is no motive to perform an act of promise-keeping that does not involve the *mistaken* apprehension that it is a duty. In the light of my subsequent discussion I should now dismiss these apparent conclusions. (2) becomes the claim that there is no *natural* (non-artificial) motive to perform an act of promise-keeping other than a sense of duty, from which it follows given (1) that an act of promise-keeping can not be a *natural* duty. The last sentence of the first passage, which I ignored earlier, in effect draws this conclusion, and indicates that promise-keeping must be an artificial duty resting on human conventions.

I read the third passage as parallel to the first, claiming that (7) an action can be equitable only if there is some motive to perform it



other than a sense of its equity, and (8) there is no universal motive to perform an act of equity other than its equity, from which it would follow that (9) the apprehension of an action as equitable is a *misapprehension*, and (10) the motive to perform an act as equitable involves this *misapprehension*. But (8) now becomes the claim that there is no natural (non-artificial) motive to perform an act of equity other than its equity, from which given (7) it follows only that there is no natural apprehension of an action as equitable. And this is in effect Hume's conclusion in the last sentence of the passage, "that the sense of justice and injustice is not deriv'd from nature, but arises artificially." The third passage seemed to rule out any real obligation to observe the laws of equity, but I may now conclude that it rules out only a non-artificial obligation.

And I read the second passage, which introduces the idea that promising requires the *feigned* willing of an obligation, as reinforcing the first. But I have now argued that we can explain both an interested obligation and a moral obligation to promise-keeping, without having to introduce a feigned act of will giving rise to one or other of these obligations. Only if we were to think of promises as giving rise to a natural (non-artificial) obligation, would we need to appeal to such an exercise of imagination. But on my interpretation, we may suppose that in making a promise each person invokes his interest in being trusted, and so being able to participate in the practice of promising, and thus gives rise directly to a natural obligation of interest, as well as indirectly to a moral obligation, to perform the act promised. Hume may be right to think that the common view of promising has recourse to a supposed act of will that must prove on reflection to be feigned. But it would seem that he has no need to burden his own account with such an act, and so no need to embrace an error theory. And indeed, if he could accept his own claims about the interest each person has in conforming to the social conventions, he would have no such need. The sensible knave suggests, however, that Hume must abandon those claims—with consequences that I want now to examine.

### III

1. Having explained the moral *approbation* attending merit or virtue, there remains nothing but briefly to consider our interested *obligation* to it, and to inquire whether every man, who has any regard to his own happiness and welfare, will not best find his account in the practice of every moral duty.<sup>10</sup>

Thus Hume begins part 2 of the concluding section of the *Enquiry Concerning the Principles of Morals*. I shall not compare the account of morality and justice in the *Enquiry* with that in the *Treatise*, but focus

only on the problem that Hume recognizes in accounting for the interested obligation to justice (he does not speak of 'natural obligation' in the *Enquiry*). Hume first canvasses the other virtues, corresponding to the natural virtues of the *Treatise* (although again he eschews the word 'natural'), and concludes:

Treating vice with the greatest candour, and making it all possible concessions, we must acknowledge, that there is not, in any instance, the smallest pretext for giving it the preference above virtue, with a view to self-interest; except, perhaps, in the case of justice, where a man, taking things in a certain light, may often seem to be a loser by his integrity. (E 282, emphasis added)

The key sentence now follows:

And though it is allowed, that, without a regard to property, no society could subsist; yet, according to the imperfect way in which human affairs are conducted, a sensible knave, in particular incidents, may think, that an act of iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy. (E 282)<sup>11</sup>

I have argued that in the *Treatise*, Hume claims that as society increases, our short-sightedness leads us not to see the interest we have in maintaining social order, but that the real interest remains. Does he express this same view in the *Enquiry*? Is the sensible knave merely short-sighted? Or is he right to think

that *honesty is the best policy*, may be a good general rule; but is liable to many exceptions: And he, ... conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions. (E 282-83)

Hume does not accuse the sensible knave of mistaking his interest. He confesses that "if a man think, that this reasoning much requires an answer, it will be a little difficult to find any, which will to him appear satisfactory and convincing" (E 283). He appeals to the heart, to the "consciousness of integrity ... cherished and cultivated by every honest man" (E 283). The knave will find that he has "sacrificed the invaluable enjoyment of a character, ... for the acquisition of worthless toys and gewgaws" (E 283). At the end of the *Treatise*, Hume offers similar remarks about the "peace and inward satisfaction" that depend upon

“strict observance of [the social virtues],” but remarks immediately “I forbear insisting on this subject. Such reflexions require a work a-part” (T 620)—of the painter, not the anatomist. But in the *Treatise* these remarks are not offered in place of an argument; the anatomist has already insisted upon the interest each person has in justice. In the *Enquiry* the sensible knave seems to have silenced the anatomist, so that only the painter remains to represent the charms of justice. But if—to continue the metaphor—the anatomical basis of the painting is unsound, we must suspect that the charms are spurious.

The sensible knave need not underestimate the interest each of us has in the social order and prosperity realized through the rule of justice. Nor need he deny the peace and inward satisfaction that some persons—but, he will emphasize, only some—derive from viewing themselves as just. But he will insist that the latter satisfaction need not suffice to ensure the former. Grant that overall the costs of unjust conduct exceed the benefits; nevertheless in many situations an individual, calculating in terms of her own costs and benefits, may expect to gain more from her own injustice than she will lose from the minuscule breach she will thereby make in the social confederacy. Thus the knave denies that each person must expect every choice that she makes between conforming to and violating the rules of justice to have an effect on the behaviour of others with consequences for her own advantage sufficient to afford her with a normally adequate motive for conformity. And as I argued previously, if this be denied then Hume’s argument for our interested obligation to justice fails.<sup>12</sup>

The sensible knave is surely right. It is utterly implausible to suppose that those who violate the requirements of justice are always led by short-sightedness to pursue a merely apparent benefit at the real expense of undermining social order. Hume is no doubt correct to insist that persons are short-sighted, and do frequently overlook the longer-term effects of their unjust conduct, but his claim “that disorder and confusion follow upon every breach of these rules [of justice]” (T 499) is simply preposterous. Whatever his view in the *Treatise*, Hume in the *Enquiry* knows better—and so appeals to the “inward peace of mind” and “consciousness of integrity” of the just person (E 283), rather than the effect of his conduct on the stability of society.

But why should the just person be conscious of his own integrity? If the sensible knave is right, and we have no interested obligation to just conduct, then why should we suppose our integrity violated or compromised by injustice? Once we recognize what Hume wishes to deny but the sensible knave rightly asserts, that moral approbation and interested obligation diverge in the case of justice, then why should our sense of integrity attach itself to our approbation rather than our interest? If Hume’s argument fails, then his rhetoric must also fail:

what hopes can we ever have of engaging mankind to a practice [such as justice], which we confess full of austerity and rigour? Or what theory of morality can ever serve any useful purpose, unless it can show, by a particular detail, that all the duties, which it recommends, are also the true interest of each individual? (E 280)

With Hume's own questions the sensible knave rests his case.

The appearance of the sensible knave in the *Enquiry* parallels the appearance of the Foole in Hobbes' *Leviathan*. In Hume's *Treatise*, and in Hobbes' earlier political writings, the interested obligation to justice is presented as unproblematic. In their earlier writings both Hobbes and Hume make an easy transition from the interest each person has in maintaining the rules of justice to an alleged interest in performing the acts required by those rules. But in their later writings both come to recognize the problematic nature of this transition. Grant that if everyone acts unjustly, then each person is worse off than if everyone were to act justly; grant also that if anyone acts unjustly, then the overall effect is worse than if he were to act justly. Nevertheless, what the Foole and the sensible knave assert is that if someone acts unjustly then frequently she is better off than if he were to act justly.

Hume recognizes that individual acts of justice, taken strictly in themselves, need not be beneficial whether to the agent or to society. He argues that taken as part of a practice of just behaviour, such acts do yield overall social benefit. What the sensible knave reminds him, as the Foole reminds Hobbes, is that a particular act taken as part of a practice may benefit society, and the practice taken as a whole may benefit each member of society, without the particular act benefitting each member, and in particular, without the particular act benefitting the agent.

2. Does Hume have a reply to the knave? Consider once again Haakonssen's interpretation of Hume's account of the sense of duty, which I shall quote here as Haakonssen reformulates it:

certain actions done out of a morally *neutral* motive (self-interest) have on the whole such good consequences and seem so clearly aimed at those consequences, that men naturally come to imagine that there is a specific motive for the actions which directs them towards those consequences. They naturally come to approve of this non-existing motive and to hate themselves for not having it. And this self-hatred is the magic formula, for it constitutes the real moral motive to do justice and fulfil promises, viz. our sense of duty.<sup>13</sup>

Suppose that we complicate Haakonssen's interpretation by recognizing that the initial self-interested motive fails to survive critical reflection, so that we come to recognize, with the sensible knave, that self-interest does not always give us a sufficient motive to act justly, or keep our promises, even when we consider our particular acts in the context of social practices. Nevertheless each person recognizes the overall good consequences of acts of justice or fidelity in the context of the relevant practices, and the benefits to each person, herself included, of the practices. And it is this recognition, coupled with the failure of self-interest to motivate her, that leads her to imagine that there is a natural motive for acts of justice and fidelity, which involves the willing of an obligation. This imagined motive receives her moral approbation. But since an obligation can not be willed, there can be no such motive, and so she comes to hate herself for not having it. This self-hatred now creates a real, genuinely existent motive to justice and fidelity—the sense of duty. The motive of duty is real, but it rests on the mistaken apprehension that we could, but do not, have a natural motive obliging us to justice.

If Hume were to answer the sensible knave in this way, then morality would not prove theoretically dispensable in motivating conformity with the artificial virtues.<sup>14</sup> It would replace the motive of self-interest in those contexts in which that motive would yield a sub-optimal outcome. But it would itself be a motive of a quite different kind. And although this motive might naturally be relatively weak, the value that each person would attach to its manifestation in his fellows would provide support for a new set of artifices—public and private indoctrination—to maintain and strengthen the initial artifices of justice, fidelity, and allegiance.

Each person would, as the knave recognizes, “be a loser by his integrity” (E 282), but each would be a gainer by the integrity of others, and would expect to gain more than his own integrity would cost. Thus each would welcome the replacement of self-interest by integrity, and would want his fellows to be scrupulous in carrying out their moral duties, even at the price of being swept up himself in moral enthusiasm. Each would recognize the value of “the artifice of politicians, who ... have endeavour'd to produce an esteem for justice, and an abhorrence of injustice” (T 500). But as Hume insists, “publick praise and blame” as well as “private education and instruction” can not create the distinction between justice and injustice (T 500). Only the peculiar working of the affections, in which the want of an imaginary motive proves to be the ground of a real one, can provide the materials needed for indoctrination to motivate us to be just.<sup>15</sup>

3. We should now return again to the passages with which I began. If we take the message of the sensible knave seriously, then it should seem that my initial reading was indeed the correct one. I read the first passage as involving the claim that (1) an action can be a duty only if there is some motive to perform it other than a sense of duty, and (2) there is no motive to perform an act of promise-keeping other than a sense of duty. Hume supposes that we feign such a motive in the willing of an obligation, and grounds our sense of duty on our finding ourselves without this imaginary motive. But since there is no real motive, it follows that both (3) an act of promise-keeping can not be a duty, and (4) there is no motive to perform an act of promise-keeping that does not involve the *mistaken* apprehension that it is a duty. The final sentence of the first passage offers no way out of these conclusions. Rather, Hume's real claim must be that an appeal to human conventions enables us to relate particular acts of promise-keeping to a beneficial practice, and so, through the peculiar workings of our affections, to be led to the mistaken apprehension that makes these acts the objects of our sense of duty. And similarly, Hume's real claim in the third passage must be that through conventions we relate particular acts of equity to a beneficial practice, and so again through mistaken apprehension we make them objects of our sense of duty.

This reading is supported by the second passage. I took Hume to claim that (5) there is no natural obligation to perform an act of promise-keeping, but (6) we feign the willing of such an obligation. The message of the sensible knave is that there is no way to bypass this feigning by appealing to a prior interested obligation. Agreeing with him requires us to dismiss Hume's reference to "a new obligation" in the first sentence. There is only the feigned obligation, and the mistaken apprehension that we might be, but are not, motivated to will it. And, hating ourselves for this supposed motivational failure, we come to regard the obligation as required by our sense of duty. If there is no interested obligation to equity or fidelity, then it would seem that Hume is able to account for the supposed moral obligation only in terms of an error theory.

Is an error theory also needed to account for the moral *approbation* that we give to a just disposition? It may seem not. For such a disposition is evidently of benefit, if not to the agent yet to his fellows, and so would seem to pass Hume's utility standard for moral approval. The idea of a just or equitable disposition pleases me *after a certain manner*; this pleasure in turn leads me to judge that the disposition is a virtue, and that its possessor is virtuous. The idea of a disposition to fidelity, or to allegiance, pleases in a similar way, and again leads to the judgements that the disposition is a virtue and its possessor virtuous. But this is only part of the story. For the sensible knave

reappears to insist that a just disposition is positively disadvantageous to its possessor, in motivating her to act justly in those situations in which she should be taking advantage of the exceptions. And if we come to regard the disposition in this light, then the idea of it will displease, leading to the judgement that it is vice.

Hume never faces the possibility that the same disposition may benefit others but be costly to its possessor—or alternatively, benefit its possessor but be costly to others. His account of moral approbation does not seem to leave room for such a possibility. I shall not speculate about the revisions that Hume would need to make in that account, were he to take the possibility seriously. I want rather to consider how Hume would account for the existence of dispositions to justice or fidelity or allegiance, if he were to accept the message of the sensible knave. For a person is disposed to justice only insofar as she takes her disposition to be a virtue and its exercise morally obligatory. Recall Hume's insistence, in the passage I quoted at the beginning of this paper, that, "If we thought, that promises had no moral obligation, we never shou'd feel any inclination to observe them" (T 518). But I have argued that in the absence of any interested obligation to justice, this moral obligation must be feigned. In feigning it we bypass the problem that the sensible knave seeks to make us face when he insists that a disposition to justice or fidelity or allegiance is beneficial to others but costly to its possessor. The disposition receives our moral approbation, but only because we have already fallen into the error of supposing ourselves under an obligation to act in accordance with its dictates.

Under pressure of the sensible knave, Hume's account of the artificial virtues becomes an error theory. Only by feigning a natural motive to the performance of just acts, do we develop the disposition to be just. But on reflection we recognize that there can be no such motive. "No action can be requir'd of us as our duty, unless there be implanted in human nature some actuating passion or motive, capable of producing the action" (T 518). Since there can be no such motive, then justice can not be required of us as our duty. Nor can fidelity and allegiance. The sensible knave's message is that human society, which depends on these dispositions, lacks any moral foundation.

4. Before the sensible knave appears on the scene, Hume asks of his account of morality:

But what philosophical truths can be more advantageous to society, than those here delivered, which represent virtue in all her genuine and most engaging charms, and make us approach her with ease, familiarity, and affection? (E 279)

## ARTIFICIAL VIRTUES AND THE SENSIBLE KNAVE

Is this the true effect of Hume's account, once its answer to the knave is shown to be an error theory? It would seem rather that Hume condemns both the effects of his theory and his delivering it, for he says,

And though the philosophical truth of any proposition by no means depends on its tendency to promote the interests of society; yet a man has but a bad grace, who delivers a theory, however true, which, he must confess, leads to a practice dangerous and pernicious. (E 279)

Should we then conclude with Hume that "mankind will agree, if they cannot refute [these pernicious truths], to sink them, at least, in eternal silence and oblivion" (E 279)?

The sensible knave has a sombre message for Hume. Does he have such a message for us? Does he reveal a problem in our own understanding of morality, or at least of justice and fidelity? Very briefly, the problem the knave reveals for Hume is this. Moral approbation is directed at dispositions because they are agreeable or useful, to their possessor or to others, and specifically moral motivation reinforces or in some cases supplies the place of these dispositions. Thus generosity is a disposition that receives moral approbation, and in virtue of her approbation, a person's natural generosity is reinforced, or her want of generosity to some extent overcome. But were there no natural disposition to generosity, then there would be nothing to approve, reinforce, or substitute for. There is no natural disposition to the particular acts we call just, but if these acts be situated within a general practice, then Hume claims that redirected interest provides such a disposition; it receives moral approbation and is reinforced or substituted for by the motive of duty. But the knave rejects Hume's claim; there is no non-moral disposition to justice, and so nothing for morality to approve, reinforce, or substitute for. Assume that we agree with the knave; he creates a problem for us only if we suppose with Hume that morality has only a second-order motivational role. Why should we suppose this?

I shall not try to answer this question here. Put more broadly, the question asks whether morality is in itself among our fundamental concerns, or whether it relates to the realization of other non-moral concerns. Some would find no threat to morality, or more specifically to justice, in the absence of an interested obligation to conform to just practices. Others would see the absence of such an obligation as indicating a deep conflict between justice and deliberative rationality. To continue the discussion would take us far beyond an interpretation of Hume's account of the artificial virtues of justice, fidelity, and allegiance.

5. But there is one feature of Hume's account that should be canvassed before I conclude, since it suggests an alternative answer to the knave that would offer less destructive consequences for our understanding of justice. Recall that Hume introduces two quite different interests that motivate promise-keeping or fidelity: the interest each of us has in the practice of promising, and the interest each of us has in himself being trusted. It seems clear that an interest of the former kind, however strong it may be, will be insufficient in itself to motivate normal compliance with the rules of justice, fidelity, and allegiance. The sensible knave must be right to claim that in many particular situations, the benefits to the individual of non-compliance will be correctly judged by him to outweigh the cost *to him* of non-compliance.

What about an interest of the second kind? Although Hume introduces it only to support fidelity, we may relate it more generally to the artificial virtues. Each person has a strong interest in being admitted into the benefits of mutual association that justice and fidelity make possible. But anyone who is not disposed to comply with the rules of justice, and to fulfil his promises and assurances, must be viewed by his fellows as an unwelcome participant in society, willing to take advantage of the just and trustworthy conduct of others rather than to contribute to mutual benefit. Such a person undermines rather than strengthens the bonds of association. Others will seek his exclusion, collectively through ostracism where this is feasible, and individually by refusing to enter into particular relations of trust with him. Thus the interest each person has in being a participant in society will on reflection lead him to be disposed to the artificial virtues, insofar as his fellows are led by their own interests to treat such a disposition as a condition of participation. Being disposed to equity and fidelity, he will be led in particular situations to act in ways that are contrary to his advantage or interest. But he will expect these real costs to be outweighed by the benefits he receives in the form of opportunities that were he not disposed to the artificial virtues, he would not enjoy. He will of course recognize that in any given situation he may expect to do less well than someone who, like the sensible knave, "takes advantage of all the exceptions" (E 283), and so seeks her own interest at the expense of justice. But he will nevertheless expect the knave to be the real loser, in forfeiting the opportunity to enter into relationships with others that depend on the expectation of fair dealing and trust.

Were persons transparent, so that their dispositions were evident to their fellows, then the argument that I have just sketched would be conclusive, and interest would direct us to embrace the artificial virtues without further ado. Hume would have a reply to the sensible knave that would not lead him to an error theory. Indeed I could simply return to the reading of Hume's account of obligation that I offered in section

II. The interest that each person would have in disposing himself to justice, fidelity, and allegiance would be at the heart of Hume's claim:

There is no passion, ... capable of controlling the interested affection, but the very affection itself, by an alteration of its direction. Now this alteration must necessarily take place upon the least reflection; since 'tis evident, that the passion is much better satisfy'd by its restraint, than by its liberty.  
(T 492)

But we should complete the passage not as Hume does (quoted above, section II.1), but by adding:

and that in making ourselves welcome participants in society, we make much greater advances in the acquiring of possessions, than in the solitary and forlorn condition, to which we should find ourselves left by our fellows were we to exhibit an universal licence.

In the real world persons are not transparent, and so whether Hume can silence the sensible knave is not proven. Since to continue the discussion would take us far beyond an interpretation of Hume's account of the artificial virtues and their obligation, it seems fitting to conclude with this Scots verdict.

*University of Pittsburgh*

1. The original version of this paper was written while a Visiting Fellow in the History of Ideas program, Research School of Social Sciences, The Australian National University, and presented to a symposium on Hume at the Flinders University of South Australia. I am grateful both to my hosts and to my discussants. I am also grateful to my colleague, Annette Baier, for her extensive comments on the second draft of the paper.
2. David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2d ed., rev., ed. P. H. Nidditch (Oxford, 1987), 518-19 (hereafter cited as "T").
3. Although Hume treats all of the artificial virtues in that part of the *Treatise* entitled "Of Justice and Injustice," he does not normally include the virtues of allegiance and chastity within the scope of justice, which is set by the three laws of nature. Thus justice includes the virtue of fidelity, or promise-keeping. But since promise-keeping is singled out by Hume for special attention, it

- will be convenient to note its distinctness by speaking at times of the virtues of justice and fidelity, or equity and fidelity, where equity is narrowly conceived as observance of the first law prescribing stability of possession.
4. The claim that Hume is an error theorist may be found in J. L. Mackie, *Hume's Moral Theory* (London, 1980), 72.
  5. Annette Baier refers to Hume's "puzzling departures in his terminology," and admits, "It is hard to save all these textual phenomena," in discussing Hume's view of natural obligation in "Promises, Promises, Promises." See Annette C. Baier, *Postures of the Mind: Essays on Mind and Morals* (Minneapolis, 1985), 183, 205 n. 19.
  6. See the discussion of this question in Baier (above, n. 5), 188-93.
  7. Bernard Wand, "Hume's Account of Obligation," in *Hume: A Collection of Critical Essays*, ed. V. C. Chappell (Garden City, New York, 1966), 326. (First published in *Philosophical Quarterly* 6 [1956].)
  8. Knud Haakonssen, *The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith* (Cambridge, 1981), 34.
  9. *Ibid.*, 35.
  10. David Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L. A. Selby-Bigge, 3d ed., rev., ed. P. H. Nidditch (Oxford, 1975), 278 (hereafter cited as "E").
  11. Annette Baier questions whether the interested obligation to justice which the sensible knave challenges in the *Enquiry* embraces the performance of promises, or instead is restricted to a concern with stability of possession. The answer is surely found in this passage, in which the knave envisages the advantages of "an act of iniquity or infidelity" (emphasis added).
  12. Barry Stroud offers a careful discussion of Hume's argument, reaching the conclusion, "If there are any occasions on which a man reasonably thinks he can get away with injustice and profit from it then there will be nothing to recommend justice to him at all" (p. 210), in *Hume* (London, 1977), 202-14.
  13. Haakonssen (above, n. 8), 35.
  14. But as Annette Baier points out, if Hume were to answer the sensible knave in this way, emphasizing the role of self-hatred, he might be hard pressed to maintain his insistence that our moral sense "reflecting on itself, ... approves of those principles, from whence it is deriv'd, and finds nothing but what is great and good in its rise and origin" (T 619). But in the *Treatise*, where this insistence appears, Hume does not acknowledge that moral approbation and interested obligation may diverge. We should not

be surprised to find that Hume's benign account of our moral sense could not fully survive this acknowledgement, and that darker passions come to the fore when our moral sense is at odds with our interest.

15. Thus Marcia Baron, in "Hume's Noble Lie: An Account of His Artificial Virtues," *Canadian Journal of Philosophy* 12 (1982) is mistaken to claim that, "The people ... must be lied to" (p. 555). She supposes that the indoctrination required to motivate us to be just depends on lying (nobly) to the effect "the benefits of justice are greater than in fact ... they are" (p. 552), so that invariable compliance with the rules of justice is made out to be in each person's interest. If I am right, each person, by imagining that there is a natural motive for acts of justice, which involves the willing of an obligation, tells herself the lie necessary to give indoctrination a foothold. The people need not be lied to, because they lie to themselves.