



### **Hume and the Prince of Thieves**

Jennifer Welchman

*Hume Studies* Volume 34, Number 1, (2008) pp. 3–19.

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 and the Prince of Thieves

JENNIFER WELCHMAN

*Abstract:* Hume's readers love to hate the Sensible Knave. But hating the Knave is like hating a messenger with bad tidings. The message is that there is a gap, on Hume's account, between our motivations and our obligations to just action. But it isn't the Knave's character that is to blame, for the same gap will be found if we turn our attention to alter egos, such as Robin Hood, the benevolent "Prince of Thieves." Replacing self-interest with benevolence not only does not make the gap go away, it makes it harder to bridge. Of the two, it is benevolence, not self-interest, that actually poses the more serious challenge to Hume's account of justice.

Lythe and listin, gentilmen,  
That be of frebore blode:  
I shall you tel of a gode yeman,  
His name was Robyn Hode.

\* \* \*

For he was a good outlawe,  
And dyde pore men moch god.<sup>1</sup>

Hume's readers love to hate the "sensible knave."<sup>2</sup> But hating the knave is like hating a messenger with bad tidings. The message Hume's sensible knave delivers is that neither our natural nor our moral obligations to justice can be relied upon to warrant our acting justly in each and every case that may arise. If this is cause

---

Jennifer Welchman is an Associate Professor in the Department of Philosophy, The University of Alberta, Edmonton, Alberta, T6G 2E7, Canada. E-mail: jennifer.welchman@ualberta.ca.

for concern, our concern should be focused upon the character of the message, not of the messenger. But too often the knave, or rather the knave's self-interest, is confusedly held responsible for the bad tidings he delivers.<sup>3</sup> Nothing else would appear to explain why so much ink has been spilt over Hume's discussion of the sensible knave and little or none on his discussion of otherwise motivated counterparts, such as Cyrus the benevolent rule breaker, though Hume uses both kinds of characters to illustrate one and the same problem, and in one and the same text.<sup>4</sup> In the following, I argue that (1) the knave's self-interest has nothing in particular to do with the problem Hume uses him to illustrate and (2) all in all, it is benevolence rather than self-interest that poses Hume's account the greater threat.

Hume introduces the infamous sensible knave as follows:

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. That *honesty is the best policy*, may be a good general rule; but is liable to many exceptions: And he, it may, perhaps, be thought, conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions. (EPM 9.22; SBN 282–33)

So the knave is a rational, self-interested individual who recognizes that while habitual obedience to his society's system of justice generally serves his self-interest best, occasions arise when he can disobey the system's rules without undermining it. Say, for example, he finds himself in a situation where he could gain substantially by a theft or fraud that will (1) go undetected and (2) neither harm nor alarm any particular persons whose subsequent distress might affect his own peace of mind.

What problem for Hume's account of justice does the knave illustrate? It is the problem Hume faces in trying to defend his claim that we are obligated to fulfill the requirements of justice whatever the circumstances, a claim that seems insupportable on his account of our moral psychology. On Hume's account it does not appear to be possible that we could be obliged to act as justice requires whatever the circumstances. So if "ought" implies "can," Hume ought to deny what he specifically asserts: that our obligation to be just is inflexible and does not vary with variations in the context of action. A brief review of three crucial components of Hume's account is necessary to show how the problem arises.

First, for Hume, our moral obligations are just that sub-set of our non-moral motivations to action that happen to be endorsed or approved when we reflect

upon them from a generally disinterested point of view. Hume writes in the *Treatise*<sup>5</sup>:

All morality depends upon our sentiments; and when an 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 3.2.5.4; SBN 517)

What makes these sentiments moral is that they are common to all normal adult human beings who adopt a disinterested point of view on the acts or dispositions under review. By contrast, sentiments arising from considering the effects of acts or dispositions upon our own interests or the interests of those to whom we are partial are not necessarily common to all who might reflect upon them and so are prudential rather than moral. “One man’s ambition,” Hume notes, “is not another’s ambition; nor will the same event or object satisfy both” (EPM 9.6; SBN 272). That an act or object frustrates a person’s ambitions or satisfies her tastes, determines that it is good or bad to her, but does not entail that it is good or bad from a disinterested or moral perspective. To be good or bad morally, it must be the case that any normal, disinterested individual would approve or disapprove it from sympathetic consideration of its effects impartially considered.<sup>6</sup> As “the humanity of one man is the humanity of every one,” Hume holds, the disinterested sympathetic response of any one person will be the same as another’s (EPM 9.6; SBN 272). Dispositions, motivations, or obligations to action that fail to elicit such approval or disapproval are not, properly speaking, moral. Regarding our obligation to justice then, if we are morally obliged to be just, it must be the case that in addition to any personal interests we have in the performance of just actions, we are also moved to approve their performance by our disinterested sympathetic appreciation of their tendency to promote others’ welfare.

Second, in the *Treatise*, Hume also declares that “no action can be virtuous, or morally good, unless there be in human nature some motive to produce it, distinct from the sense of its morality” (T 3.2.1.7; SBN 479). That is, dispositions capable of motivating action cannot be created *ex nihilo*. Nor can they be created by speculative reasoning to the effect that a particular disposition would be morally approved if, contrary to fact, we possessed it, because “reason is perfectly inert, and can never either prevent or produce any action or affection” (T 3.1.1.8; SBN 458). Thus to be morally obliged to act justly, we must already have been disposed to act justly by a disposition or dispositions we possessed or acquired prior to and independent of subsequent reflection from the moral point of view. In the case of justice then, if we are to be morally motivated to just action, it must be the case that

we are already non-morally motivated to just action in ways that we subsequently approve on reflection from the disinterested, moral point of view.

Third, Hume also holds that the disposition to act justly is not an innate or “natural” human endowment, but rather a disposition acquired through education, training, and other forms of social artifice, for which reason he calls justice an “artificial” virtue.<sup>7</sup> Human beings are habituated from childhood to conform to the conventional system of justice into which they are born. Once the habit of conformity to a social convention is acquired, the process by which mature individuals come to morally approve this acquired disposition, and acts in accordance with it as obligatory, follows the same pattern. In the case of justice, we first approve our habitual conformity to conventions of justice, prudentially, from narrow self-interest, and subsequently approve it morally as we reflect upon it from a more disinterested point of view. We note that some of those whom the system of justice protects are our friends and family, to whose well-being we are partial. Then we go on to reflect that it also serves others for whom we have no strong partiality, but to whose joy and suffering our sympathy makes us sensitive, with the result that we deplore attacks upon their persons and properties for their own sakes as well as our own. Thus we come to see acts that preserve or enhance the system as *morally virtuous* and acts undermining it as *morally vicious*. As we are displeased in the relevant manner by non-performance of acts required by our conventional systems of justice, performing them is then obligatory.

Artificial virtues like justice are approved by us for their personal and social utility. However, as Hume himself points out in both the *Treatise* and the second *Enquiry*, situations can arise where the consequences of acting as justice requires can undermine our approval for specific instances of just action. All actual conventions of justice are complex systems of rules and requirements, a substantial percentage of which are, strictly speaking, superfluous for maintaining that system or realizing its goals of personal and social utility. Since our approval of justice is grounded in its utility, whenever we are faced with one of these unnecessary rules or requirements, our disapproval of non-performance of those specific requirements will be weakened. Worse yet, in some cases, compliance entails consequences that are directly personally or socially pernicious. Hume notes, any system of justice may “deprive, without scruple, a beneficent man of all his possessions, if acquired by mistake, without a good title; in order to bestow them on a selfish miser, who has already heaped up immense stores of superfluous riches” (EPM App 3.6; SBN 305–306). He continues, “And though such rules are adopted as best serve the same end of public utility, it is impossible for them to prevent all particular hardships, or make beneficial consequences result from every case.” In these cases, Hume grants we must expect that a person’s “benevolence and humanity, as well as his self-love, might often prescribe to him measures of conduct very different from those, which are agreeable to the strict rules of right and justice” (EPM App 3.7;

SBN 306). Though we will continue to approve of general conformity to conventions of justice, we are apt to approve rather than disapprove of non-conformity with justice in those cases where conformity is patently pernicious to ourselves or others. No doubt the force of habit will press us to resist following the prescriptions of benevolence, humanity, or self-love nonetheless. But once we have recognized that the force of habit is all that obliges us to justice in such cases, we cannot help but ask ourselves if we would not be fools to allow the discomfort of overriding the habit to outweigh the pleasure we would get from indulging our self-interest or benevolence. “Who . . . feels not a secret sting or compunction, whenever his memory presents any past occurrence, where he behaved with stupidity or ill manners?”—Hume asks (EPM App 4.3; SBN 314–15.) No one wants to play the fool or to be humiliated, even if only in their own eyes. Thus, if we become convinced it would be mere foolishness to act as justice requires in a particular case, our pride will provide a further inducement to resist the force of the habit of justice.

Hume takes pains to convince us that acting as habit directs in such cases is practically rational in the long run, by stressing the risks of over-riding our habits of just action. We may be caught and punished. And if we escape punishment, our escape may encourage others to follow our example, thereby weakening the common commitment to social justice and threatening the peace and security it affords us. But while Hume’s arguments provide us with sufficient reason to allow our habits to direct our conduct in most such cases, they seem to fall short of providing sufficient reason for believing we should allow them free rein on every occasion.<sup>8</sup>

So if Hume means to provide us sufficient reason to view just action as obligatory in every possible case, he does have a problem here—a gap between what he holds our obligations to be and his psychological account of how they arise. It is, however, not a problem that has any particular connection to self interest or to the moral psychology of the sensible knave. Yet so great is the human tendency to blame the messenger for the message he brings, these considerations alone are unlikely to be sufficient to dispel the impression that nevertheless human self-interest is somehow peculiarly to blame. Possibly the only way this can be done is by showing that the message would be just as bad, and quite possibly worse, if the messenger bringing it was a paragon of benevolence. To this end, I beg your indulgence of the following philosophical fable: *Robin Hood’s Progress*.

**Prologue:** Forget for a moment the many Hollywood biopics of Robin Hood. Our Robin is not a dispossessed member of the nobility. He is the outlaw of medieval legend, a simple yeoman who has been forced to flee his holdings and start life anew in another part of the country.<sup>9</sup> With only a few groats in his pocket, Robin takes refuge in a forest, to live by his wits and his bow until something better comes along. Shortly after his arrival, Robin finds the bodies of a knight, a royal tax collec-

tor, and his horse, drowned, on the bank of a swift-moving river they had unwisely attempted to ford. Miraculously, the knight's bags have not all been washed away. One lies on the bank, bursting with tax money collected on behalf of the crown. Though well brought up and habituated to just action, the sight of so much money shocks Robin out of his customary patterns of thought and behavior. He realizes he is alone and unobserved. If he took the money and threw the bag into the river, no one would guess that the knight had been robbed. Justice demands he turn the money in to the local authorities. Is he obliged to comply?

### Fit I: The Knave

If Robin is a genuinely sensible knave, say a knave out of Bernard Mandeville's *Fable of the Bees*, then the only passion strong enough to control his self-interest will be self-interest itself. He must be persuaded that his self-interest is better served by returning the money than by keeping it. Being sensible, he judges that very probably it is. He is a stranger in these parts, a person no one has any particular reason to trust or assist. How can he earn the locals' good will, so essential to his long-term welfare? A Mandevillian<sup>10</sup> will reason as follows:

The Man of Manners picks not the best but rather takes the worst out of the Dish, and gets of every thing, unless it be forc'd upon him, always the most indifferent Share. By this Civility the Best remains for others, which being a Compliment to all that are present, every Body is pleas'd with it: The more they love themselves, the more they are forc'd to approve of his Behaviour, and Gratitude stepping in, they are oblig'd almost whether they will or not, to think favourably of him. After this manner it is that the well-bred Man insinuates himself in the esteem of all the Companies he comes in.<sup>11</sup>

In the short-run, stealing the money has greater personal utility than turning it in to the local authorities. But in the long-run, the probabilities go the other way. By giving up the money, the knave will instantly make himself a reputation for honesty, justice, modesty, and loyalty ("No really, m'lud, I only did what any decent man would do") that will serve his interests far better in the long run. He will be clapped on the back, hailed as a "good fellow," stood drinks in the local pub, and befriended by one and all, most especially the local representatives of the crown, the people best positioned to advance his fortunes in that community.<sup>12</sup> As he arrives at this conclusion, Robin's pride is tickled by the thought of how easily he can manipulate the situation to his own advantage.

## Fit II: The Self-centered Humean

Then Robin laughs at his own thoughts. He knows he is not really the Mandevillean knave just described. He is simply a particularly self-interested version of the sort of creature Hume takes us all to be, attuned by sympathy to one another's pleasures and pains, and naturally disposed to act for others' welfare by dispositions such as love and (limited) benevolence. He will of course give up the money, as it is clearly in his long-term material best interest to do so. But the pride he takes in his action will have a different source than the Mandevillean knave's. A Mandevillean's sense of self-worth would presumably be based solely upon his possession of dispositions or skills that maximize his *own* welfare<sup>13</sup>—as these are the only traits he could assume other sensible knaves would envy or admire him for possessing. A Humean by contrast, who supposes human beings generally possess other-regarding motivations as well as self-regarding ones, can expect to be admired for possessing and acting from dispositions that do himself no direct or immediate good so long as they tend to promote the welfare of others. Thus, his pride can be tickled by the thought that giving up the money is evidence that he actually possesses the habits of just action that he knows others generally admire and approve.

Robin slings the knight's bag over his shoulder and begins the journey into the town, his burden lightened by the thoughts that (1) very probably, he will enjoy all the indirect benefits he anticipated when he reviewed the situation from the knave's perspective and that (2) even if his hopes are in some measure disappointed, this will be counterbalanced by the pleasure he can take in knowing himself to possess a disposition that warrants pride. Of course, if fulfilling the demands of justice directly and immediately threatened his own life or limb, his self-interested obligation to justice would be defeated and so too his moral obligation. But this is not, I think, an exception to the rules of justice that either Hume or any reasonable society should scruple to accept.

## Fit III: The Lover

The path to the village passes the hovel of a poor, malnourished woodcutter and his pretty daughter Marian—the only people in the district with whom Robin has as yet become acquainted. Robin likes the woodcutter and fancies himself in love with Marian. Sadly, both suffer from the ague. Unless their diet, clothing, and shelter swiftly improve, both may die. Robin halts. If he were to pocket the tax money and dispose of the bag, he could help Marian and her father. He asks himself whether he really ought not steal in this case, or indeed, in others like it.

Robin's concern for his own welfare is now submerged and effectively silenced by his benevolent and loving concern for Marian and her father. Does this mean

that his habit of justice is now unopposed? Is Robin more likely to feel obliged to be just than when his personal self-interest held sway? On the contrary, his habit of justice is now opposed by his love and benevolence, each of which presses him to override the habit in the interests of Marian and her father. Unless he can be persuaded that his secret theft will do them more harm than good in the long run, he will conclude that he would not only be a fool but a callous, coldhearted blackguard, were he to neglect this opportunity to help his ailing friends. Can he be persuaded? Probably not, if he is convinced that one or both will perish without his help. If so, perhaps Hume and any reasonable society should be prepared to allow that the rules of justice may be suspended in so extreme an emergency. But realistically, Robin can only be certain that Marian and her father will suffer miserably for days or weeks, leaving open the possibility that it could be rational to trade off their short-run suffering against a greater long-run advantage to them, assuming there is any, in Robin's acting as justice requires by returning the money.

At the same time it also opens up the possibility of trading off advantages to these two individuals against advantages to others in the district. Benevolent individuals can have and act from either or both of two forms of benevolence Hume distinguishes: particular and general. Particular benevolence is a relatively narrow form of benevolence, "founded on an opinion of virtue, on services done us, or on some particular connexions" we have with those to whom we are partial. General benevolence, on the other hand, is the wider benevolence from which we act "where we have no friendship or connexion or esteem for the person, but feel only a general sympathy with him or a compassion with his pains, and a congratulation with his pleasures" (EPM App 2.5n60; SBN 297-98n60). Robin's motivation to help Marian is particular benevolence. But because he is also disposed to general benevolence, he is also concerned for the welfare of others in the district with whom he is not yet personally acquainted. Now were those others to suspect a theft has been committed and become alarmed at the thought of a thief operating in the forest, then the good his theft achieved for Marian would come at the expense of a more general harm to her neighbors. Of course, Robin has no reason to think that his crime will be detected, but perhaps he should fear that his success might embolden him to make it a rule to seek similar opportunities in future. If so, he might reasonably fear that eventually his crimes would be detected and so cause general distress.

Poor Robin is in a quandary, his two sorts of benevolence pressing him in opposing directions. He tries to settle the matter by reflecting on his contrary motivations from a disinterested, moral point of view. But when he does so, the guidance he receives is contradictory. Evaluation from the moral point of view requires him to exclude his personal interests from consideration, which in this case means that he must try to imagine how disinterested observers would feel were Robin to return the money or use it to assist Marian and her father. To a dis-

interested observer, Marian's welfare would be no more important than that of any other individual whom Robin's acts affect. However, a disinterested observer would also expect Robin to give priority to the interests of those to whom he is partial. Consequently, such an observer would be apt to blame Robin if this expectation were disappointed. Hume holds that "absence of a virtue may often be a vice; and that of the highest kind. . . . Where we expect a beauty, the disappointment gives an uneasy sensation, and produces a real deformity" (EPM 7.10n42; SBN 314–15n42). Should Robin's concern for Marian seem lacking, disinterested observers will condemn him as callous and unfeeling. As Hume notes: "we blame a person, who either centers all his affections in his family [or intimate circle], or is so regardless of them, as, in any opposition of interest, to give the preference to a stranger, or mere chance acquaintance" (T 3.2.2.8; SBN 488). Robin concludes that his only hope of winning approval from the moral point of view is to find a way to exhibit both sorts of benevolence at once.

He wonders if there is some way he can, in effect, have his cake and eat it too: commit injustice from partial benevolence while still exhibiting his more general benevolent concern to protect the welfare of others in the district. Happily two possibilities immediately occur to him. He could steal the tax money for Marian's sake and then take steps to ensure that theft does not become a habit, by resolving thereafter to leave the forest forever and settle in such close proximity to others as to make secret thievery impossible. Alternately, he might use part of his ill-gotten gain to help Marian, and then after earning back the sum he spent upon her by honest means, turn himself and the money into the authorities, denouncing his own misuse of the money and inviting the authorities to impose any penalties they felt necessary to ensure that public confidence in the justice system is upheld.

#### **Fit IV: The Princely Thief**

Something in Robin rebels at these ways of reconciling his particular and general benevolence. Although he is but a simple yeoman, Robin's character is marked by what Hume characterizes as "GREATNESS OF MIND, or Dignity of Character, with elevation of sentiment, disdain of slavery, and with that noble pride and spirit, which arises from conscious virtue" (EPM 7.4; SBN 252).<sup>15</sup> Such a man would not submit himself to continual public scrutiny for fear of immoral action. Nor would he denounce himself for doing what he thought right. If he stole a thing, he would do openly and unabashedly. So if he is to reconcile his particular and general benevolence, it must be by some other means.

He thinks to himself that times have been hard. Premature death from malnutrition is common in the district. The taxes being collected could be put to better use if distributed to all the poor boxes of the local parish churches, or better

yet, given directly to the poor, than if added to royal coffers to pay for yet more foreign wars. His theft of the tax money would, if detected, cause some disquiet, no doubt. Yet that same disquiet might have beneficial consequences. He recalls having heard of a bandit, also called Robin, whose notoriety had just such an effect: “*So that through dread of Robbin then, and his adventurous crew, the mizers kept great store of men, which else maintayn’d but few.*”<sup>16</sup>

Suppose he were to make it publicly known that he would henceforth steal from the rich and give to the poor, unless or until the rich volunteered to support the destitute of the district. This, it seems to him, is surely how a man who possessed greatness of mind and general benevolence would behave.<sup>17</sup> Robin knows that if he does this, the local sheriff will pursue him and that eventually he will probably be caught or killed. But this is a risk he is prepared to take. With his mind made up and his conscience clear at last, he pockets the money and heads into town. There he buys food for Marian and her father, distributes the remaining money amongst the local poor boxes, sends messages to like minded-friends to join him, and then retreats into the forest. A legend is born.

**Epilogue:** The moral of the fable is simply stated. Benevolence poses at least as great a challenge to Hume’s account of our obligation to justice, as does the self-interest of sensible knaves, and quite possibly a rather more serious challenge. The primarily self-interested Robins of Fits 1 and 2 are each more easily persuaded to conform to the requirements of medieval justice than the benevolent Robins of Fits 3 or 4. The more altruistic Robin becomes, the more unwilling he is to act as justice requires. This suggests that arguments designed to bolster the altruistic impulses of sensible knaves in hopes of motivating them to be just, might, if successful, only aggravate the problem they were meant to solve. Encouraging people to focus more upon their own enlightened self-interest, rather than less, may be Hume’s best option for reducing the frequency and severity of the situations where the motivational gap between our general approval for conventions of justice and for the performance of specific requirements could result in outright rebellion against the convention itself.

What might Hume say in reply? His first thought might be to try to explain away any moral approval we may feel for the princely thief Robin has become as a peculiarity of the times in which we suppose him to have lived. As Hume notes, “Among the ancients, the heros in philosophy, as well as those in war and patriotism, have a grandeur and force of sentiment, which astonishes our narrow souls, and is rashly rejected as extravagant and supernatural” (EPM 7.18; SBN 256–57). But times have changed. Although in centuries past, a princely thief like Robin Hood would perhaps have had little reason to think that the system of justice existing in his day actually benefitted the majority, and so felt no obligation from general benevolence to uphold contemporary rules of justice, this is no longer the

case. “They in their turn, I allow,” Hume writes, “would have had equal reason to consider as romantic and incredible, the degree of humanity, clemency, order, tranquility, and other social virtues, to which, in the administration of government, we have attained in modern times” (EPM 7.18; SBN 256–57). By 1751, Hume might claim, England’s system of justice had evolved into a form adequate to command the full moral approbation of his generally benevolent contemporaries. Thus an eighteenth-century Robin Hood would be in no position to claim that general benevolence would commend his roguery as “princely.”

Twenty-first century readers might not agree that eighteenth-century systems of justice were sufficiently well-formed to ensure this conclusion. Still some might argue that if not by 1751, surely in our own time, we can safely say that most currently existing Western systems of justice are adequate to command the full moral approbation of all our generally benevolent contemporaries. Whatever might have been the case in the Middle Ages, our contemporary systems of justice are such that general benevolence could never warrant overriding one’s habits of justice. Consequently, in contemporary Western societies, it is self-interest not benevolence that offers the most potent challenge to our obligation to justice.

Two rejoinders come immediately to mind. First, it might be argued that even if our existing systems of justice are so highly evolved as to make twenty-first century princely thieves unthinkable, thievery by less great-minded, more narrowly benevolent, loving thieves is not. Second, it might be argued that it is by no means clear that twenty-first century princely thieves really are unthinkable, given the evident discontent so many feel in regard to our existing systems of justice. The civil disobedience of eco-saboteurs, radical animal activists, abortion rights opponents, and a host of other morally motivated civil disobedients cannot plausibly be put down to mere self-interest. And if some of these individuals are sufficiently great-minded as to be insulated from the force of appeals either to their own self-interest or the self-interest of particular others towards whom they are partial, it is difficult to see what if anything we could appeal to in Hume’s moral psychology to persuade them to refrain from injustice. The long tradition of social admiration for the princely thieves of song and story is some indication of the real difficulty of the undertaking.

Nevertheless, it might be argued, everything said up to this point overlooks an important historical fact: Hume specifically confessed himself unable to answer the sensible knave’s self-interested ratiocinations in the second *Enquiry*. He makes no similar confession in his discussions of benevolent law breakers, such as the young Cyrus. Surely Hume would not have made this confession had he not believed that self-interest posed special problems for his account of justice that benevolence did not.

In his discussion of the sensible knave, Hume does indeed confess himself unable to answer *some* individual’s reasoning satisfactorily. But the individual in

question is not the sensible knave. He is instead any “man” in Hume’s audience who thinks that the knave’s reasoning “much requires an answer” (EPM 9.23; SBN 283). Who would such a man be? Given the nature of the reasoning he wants answered, the man in question must be a proponent of the “selfish theory of morality,” which Hume has already given his reasons for rejecting. If so, Hume’s confession here is a confession of an inability to satisfy a reader who obstinately persists in believing with Mandeville or Thomas Hobbes that human beings are thoroughgoing egoists, in the face of all the counter-arguments that Hume has already supplied. Such a reader will only be satisfied with an answer to the sensible knave grounded in precisely the sort of egoistic moral psychology that Hume has officially rejected.

Not surprisingly, at this point, Hume simply throws up his hands. If some one among his readers willfully persists in clinging to the selfish theory of human nature despite the evidence against it, he has nothing left to say except that “if his heart rebel not against such pernicious maxims, if he feel no reluctance to thoughts of villainy nor baseness, he has indeed lost a considerable motive to virtue” from Hume’s perspective (EPM 9.23; SBN 283). Indeed, he has lost any motive to virtue at all, as Hume understands virtue. It is a consequence of this man’s selfish theory of human nature that he must reject Hume’s view that disinterested moral approbation of actions is even possible, let alone capable of obliging individuals to action. So Hume concludes that “we may expect, that his practice will be answerable to his speculation” (EPM 9.23; SBN 283). That is, we may expect that a self-deluded reader of this sort will feel bound to try to act only in ways he can justify in terms of his own speculative theory of human nature. Since this theory is not Hume’s, Hume turns his attention to his more “ingenuous” readers, whose innocence of any acquaintance with or commitment to psychological egoism allows them to accept what Hume takes to be the truth about us. We are primarily self-interested. But we are also sympathetic, loving, and benevolent. We are, in other words, not merely practically rational agents, but moral agents as well.

Read this way, the “confession” offered in Hume’s discussion of the sensible knave lends no support to the practice of blaming the character of the knave for the character of the message he conveys. As *Robin Hood’s Progress* illustrates, self-interest is neither the only nor the most serious obstacle to just action. On the contrary, the more predominant the altruistic dispositions in an individual’s character, the more difficult it can become to find ways to bridge the motivational gap between that individual’s approval for conventions of justice and her or his approval of specific acts of conformity whose consequences are pernicious. And if strongly altruistic individuals are great-minded enough to disregard their own personal welfare, bridging the gap may be wholly out of the question. A Humean’s best option for reducing the frequency and severity of the situations where the

motivational gap might erupt into outright rebellion against justice is to hope for more rather than fewer sensible knaves.

## NOTES

I am very grateful for the many helpful suggestions made by anonymous reviewers and the editors of *Hume Studies*.

1 “Gest of Robyn Hode,” in *Rymes of Robyn Hood: An Introduction to the English Outlaw*, ed. R.B. Dobson and J. Taylor (Phoenix Mill, Glouc.: Sutton Publishing, 1997), 79–112. The excerpts quoted come from the first and last stanzas.

2 David Hume, *An Enquiry Concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Oxford University Press, 1998), henceforth “EPM” and cited by section and paragraph, followed by the page number in *Enquiries Concerning the Principles of Human Understanding and Concerning the Principles of Morals*, ed. L.A. Selby-Bigge, revised by P.H. Nidditch (Oxford: Clarendon Press, 1975), henceforth “SBN.” For Hume’s discussion of the sensible knave, see EPM 9.22–24 (SBN 282–83).

3 The Knave has received too many replies to mention them all here (some notable recent examples are listed below.) While none of the authors of these discussions denies that altruism as well as self-interest can conflict with justice, the relative lack of attention paid to conflicts arising from altruistic motives can seem to suggest that these are somehow less serious than conflicts arising from self-interest. See, e.g., Jason Baldwin, “Hume’s Knave and the Interests of Justice,” *Journal of the History of Philosophy* 42.3 (2004): 277–96; James King, “Pride and Hume’s Sensible Knave,” *Hume Studies* 25.1–2 (1999): 123–37; Steven Darwall, *The British Moralists and the Internal ‘Ought’: 1640–1740* (Cambridge: Cambridge University Press, 1995); David Gauthier, “Artificial Virtues and the Sensible Knave,” *Hume Studies* 18.2 (1992): 401–27; Annette C. Baier, “Artificial Virtues and the Equally Sensible Non-Knaves: A Response to Gauthier,” *Hume Studies* 18.2 (1992): 429–39; Gerald Postema, “Hume’s Reply to the Sensible Knave,” *History of Philosophy Quarterly* 5.1 (1988): 23–40; Michael J. Costa, “Why Be Just? Hume’s Response in the *Inquiry*,” *Southern Journal of Philosophy* 22.4 (1984): 469–80.

4 By contrast with the extensive literature on the sensible knave, there is essentially no literature on Hume’s discussion of this benevolent counterpart to the sensible knave, the case of the young future Emperor Cyrus I of Persia, discussed in EPM App. 3.4 (SBN 304–305), whom Hume describes as follows: “Cyrus, young and unexperienced, considered only the individual case before him, and reflected upon the limited fitness and convenience, when he assigned the long coat to the tall boy, and the short coat to the other of smaller size. His governor [tutor] instructed him better; while he pointed out more enlarged views and consequences, and informed his pupil of the general, inflexible rules, necessary to support general peace and order in society.” But Hume’s endorsement of the tutor’s lesson should not be taken as an endorsement of the view that benevolent individuals can never have adequate reason for disobedience to the letter of the law. For according to Xenophon, Cyrus’s tutor also tells him that “when I was a judge of a good fit, I should do as I had done.” Xenophon, *Cyropaedia*, 2 vols.,

trans. Walter Miller (Cambridge, MA: Harvard University Press, 1983), 1.3.17. As Hume's contemporaries would have known, what Cyrus actually learned from the incident was not that he should always obey the dictates of the Persian system of distributive justice in which he was being trained—a system that routinely assigned goods to individuals without regard to merit or social utility—but rather that he should obey them only as long as he was too immature to make sufficiently reliable judgments of personal merit or social utility to warrant his overriding the dictates of justice in particular cases. (Once Cyrus was confident that he was a good judge of what was fitting, he felt free to override traditional norms whenever he judged it appropriate.) Well informed, mature, and benevolent Humeans, like Cyrus, must inevitably draw the same conclusion—that in some cases they will have little motivation to act as justice requires and a very strong ones (such as benevolence) to do otherwise.

5 David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary Norton (Oxford: Clarendon Press, 2000), abbreviated “T” in the text and cited by Book, part, section, and paragraph number; *A Treatise of Human Nature*, ed. L.A. Selby-Bigge, 2nd ed., revised by P.H. Nidditch (Oxford: Clarendon Press, 1978), abbreviated “SBN” in the text and cited by page number.

6 Hume writes “in judging of characters, the only interest or pleasure, which appears the same to every spectator, is that of the person himself whose character is examin'd; or that of persons, who have a connexion with him . . . they alone produce that particular feeling or sentiment on which moral distinctions depend” (T 3.3.1.30; SBN 590–91). (Compare EPM 9.6–8; SBN 272–74.) Hume goes on to note that our response to others' character or conduct from the moral point of view may be too weak to override our personal subjective responses when these conflict. Still the test allows us to determine which responses we are entitled to suppose will be shared and thus moral. (See T and EPM *ibid.*) On Hume's construction of the general point of view generally, see Elizabeth S. Radcliffe, “Hume on Motivating Sentiments, the General Point of View, and the Inculcation of Morality,” *Hume Studies* 20.1 (1994): 37–58; Rachel Cohon, “The Common Point of View in Hume's Ethics,” *Philosophy and Phenomenological Research* 57.4 (1997): 827–50; and Kathleen Wallace, “Hume on Regulating Belief and Moral Sentiment,” *Hume Studies* 28.1 (2002): 83–111.

7 The term “artificial” virtue is not used as extensively in the EPM as in T, but this is no reason to suppose Hume has turned his back on this useful distinction (esp. given his comment on the subject at EPM, App. 3.9n64; SBN 307n64). In EPM, Hume prefers to distinguish virtues by differences in the sources of their moral approval as considered from the general moral point of view (i.e., utility versus immediate agreeableness) rather than by reference to differences in the manner by which we acquire them (natural endowment or education). But as these two ways of distinguishing virtues are not incompatible but rather cross-cutting, the fact that Hume opts to employ the former rather than the latter in EPM is not proof that the latter has been renounced or that commentators would be wrong to employ it when convenient.

8 Critical responses vary as to the adequacy of Hume's appeals to our self-interest to support us in our habit of just action when just action appears to be personally and/or socially pernicious. Some have argued that they are clearly inadequate and could only succeed if we are erroneously persuaded that every act of injustice will have pernicious consequences (see, e.g., Marcia Baron, “Hume's Noble Lie: An Account of His Artificial

Virtues," *Canadian Journal of Philosophy* 12.3 (1982): 539–55, and Gauthier, "Artificial Virtues and the Sensible Knave"). Others argue that while they are inadequate by themselves, Hume can appeal to other kinds of costs of injustice that will for the most part fill the gap, such as the pain of alienation from others that the necessity of concealing our injustice will force upon us (e.g., King, "Pride and Hume's Sensible Knave"), the pain of self-hatred if we see our temptation to injustice as a sign of personal weakness (see Knud Haakonssen, *The Science of a Legislator* (Cambridge: Cambridge University Press, 1981), 30–35), or the pain of frustrating our other-regarding, natural virtues, such as compassion, that make us reluctant to engage in any action that might harm the material interests of others, even if those harms were trifling in comparison to the material good that injustice offers us (e.g., Baldwin, "Hume's Knave and the Interests of Justice," and Baier, "Artificial Virtues and the Equally Sensible Non-Knaves"). Those offering such arguments acknowledge that these added inducements to justice still leave a small gap open between our obligations and motivations to be just, but argue that only rare or freakish individuals could have their obligation to justice seriously destabilized by it. Thus they argue we should not see it as a serious problem for Hume. Others are less sanguine about the prospects of bolstering our habits of justice in these ways and so from intellectual charity argue that Hume should not be read as holding that every virtuous disposition must arise prior to or independently of reflective moral approval for them as virtuous. They suggest that he should instead be read as granting that moral approval for certain rules or practices can be motivating even in the absence of any non-moral motive to that action. (See, e.g., Darwall, *British Moralists and the Internal 'Ought,'* and also his "Motive and Obligation in Hume's Ethics," *Nous* 27.4 (1993): 415–48; and Rachel Cohon, "Hume's Difficulty with the Virtue of Honesty," *Hume Studies* 23.1 (1997): 91–112.) As the object here is to explore the nature of the gap that seems to arise between the requirements of justice and our motivations to comply on standard readings of Hume, revisionist readings will not be considered. Nor will the various attempts to bridge the gap created by self-interested considerations be taken up, as the project of this paper is to show that self-interest is not the only or most important cause of the gap.

9 In the earliest surviving tale, the late medieval "Gest of Robyn Hode," Robin is a yeoman (free man and small holder) who has failed to appear in a court of law to answer a criminal charge and so been declared an 'outlaw.' He and his confederates live in the Forest of Barnsdale (Yorkshire) plundering passers by. Although he is respectful of women and willing to assist strangers, the medieval Robin's thievery was primarily self-interested. Only gradually, over several centuries, did Robin become a benevolent do-gooder who robs the rich to give to the poor. In the tale told here, Robin's centuries-long moral progress is compressed into a single morning. For the history of Robin Hood, see Dobson and Taylor, *Rymes of Robyn Hood*.

10 That is, presuming Mandevillean knaves are, as Hume depicts them, moral "monsters" whose self-interest is unalloyed with sympathy or kindness for others. This there is reason to doubt. Mandeville's moral agents more nearly resemble Hume's than they do the agents of other selfish theorists, and so a Mandevillean's deliberations might more nearly resemble those of self-centered Humeans, discussed below in Fit 2, than the morally monstrous knave discussed here in Fit 1. However, this possibility does not affect the argument offered here and so may be passed over for the present.

11 Bernard Mandeville, *The Fable of the Bees, or Private Vices, Publick Benefits*, vol. 1, ed. F. B. Kaye (Oxford: Clarendon Press, 1924), 78.

12 Note that this would seem to be the case in any comparable situation in which injustice offers great material advantages with little risk of discovery. These are precisely the situations in which publicity turns self-restraint to a knave's best social advantage. The greater the temptation to injustice is acknowledged to be, the greater the social approval and social trust extended to the man or woman known to have rejected it.

13 In this case, the Mandevillean knave can pride himself on his prudence, forethought, and strategic rationality, qualities any Mandevillean would value possessing.

14 There is considerable debate about how we are to understand Hume's almost Hutchesonian talk about benevolence in EPM in light of its apparent incompatibility with the "circles" account of benevolence in T. Possibly Hume's position on benevolence actually changed. Equally possibly, it has not; in which case what goes under the ordinary-language term "benevolence" in EPM is not just the passion of benevolence, as that is characterized in T (T 2.2.6; SBN 366–68), but also all those passions that Hume claims are routinely felt together with and as indistinguishable from benevolence, e.g., compassion and pity (T 2.2.7. and 2.2.9.4; SBN 368–72, 381–89.) In favor of the latter interpretation, we might cite the fact that in T 3.3.3.2 (SBN 602), Hume allows that our benevolence and generosity, though they grow the less intense with social distance from their potential targets, can nevertheless extend well beyond our intimate circle: "the generosity of men is very limited, and . . . seldom extends beyond their friends and family, or at most, beyond their native country" (my emphasis). We may also note that when Hume is most insistent about the relative weakness of benevolence to strangers to move us to action, he is usually speaking of situations where the countervailing motive is self-interest. See, e.g., T 3.2.2.13 (SBN 492), where Hume writes: "Tis certain, that no affection of the human mind has both a sufficient force, and a proper direction to counter-balance the love of gain, and render men fit members of society, by making them abstain from the possessions of others. Benevolence to strangers is too weak for this purpose; and as to the other passions, they rather inflame this avidity, when we observe, that the larger our possessions are, the more ability we have of gratifying all our appetites." But self-interest is not the countervailing motive for Robin in Fits 3 and 4. However, it does not matter for current purposes with which positions one sides. In this paper, general benevolence will be understood to extend no farther than Hume allows in the *Treatise*, i.e., to fellow members of one's society. For a more general discussion of the issue, see Annette C. Baier, "How Wide Is Hume's Circle? (A Question Raised by the Exchange between Erin I. Kelly and Louis E. Loeb)," *Hume Studies* 32.1 (2006):113–17.

15 For recent discussions of Hume on greatness of mind, see Kate Abramson, "Two Portraits of the Humean Agent," *Pacific Philosophical Quarterly* 83.4 (2002): 301–34, and Graham Solomon, "Hume on Greatness of Soul," *Hume Studies* 26.1 (2000): 129–42.

16 Martin Parker, "A True Tale of Robin Hood," in John Mathew Gutch, ed., *A Lytell Geste of Robin Hood*, vol. 2, (London: Longman, Brown, Green, and Longmans: 1847), 89–106, 96.

17 Robin's reflections at this point may seem to resemble those of Christine Korsgaard's "slightly more attractive version of Hume's sensible knave," a probate lawyer tempted

to suppress a rich client's most recent will, benefitting his worthless nephew, in favor of an earlier will, leaving his estate to medical research. Because the lawyer can easily suppress the latter without undermining the system of justice which she generally endorses, Korsgaard reasons that the lawyer's habitual disapproval of unjust acts may seem to her to be "in this case, poorly grounded, and therefore in a sense irrational. And this may lead her to set it aside, or if she can't, to resist its motivational force." Korsgaard suggests that she might ask herself, "since I approve of just actions because they are, generally speaking, useful, why not simply do what will be most useful?" But if she acts on this, Korsgaard argues, "she is not a Humean anymore; she is a utilitarian." See Christine Korsgaard, *The Sources of Normativity* (New York: Cambridge University Press, 1996), 86–87. So one might worry whether Robin the Princely Thief has become a utilitarian, in which case his ruminations would be irrelevant to the problem under consideration. We should note, however, that Korsgaard's conclusion about the lawyer only follows if what the lawyer means when she proposes to "do what will be most useful" is that she should henceforth suspend her support for her society's conventions of justice altogether in order to do whatever will maximize the happiness of any individuals whom her acts affect. Robin is proposing nothing of the kind. He is merely considering whether to suspend his support for particular provisions of his society's conventions of justice in light of their effects upon a specific group of his fellow countrymen and women. His general benevolence is neither as impartial nor as extensive as a utilitarian's. Moreover, whereas a utilitarian will only approve of dispositions that tend to promote others' happiness, Robin also approves of dispositions for their "immediate agreeableness" to their possessors or those about them, independent of considerations of their utility for maximizing general happiness. Take for example his approval of greatness of mind, which plays a crucial role in his deliberations in Fit 4. His approval for the course of action to which his general benevolence tempts him is partly grounded in his prior and independent approval of greatness of mind. So Robin is still quite definitely a Humean.

