

Hume on the Artificial Virtues¹

Introduction

My aim in this paper is to make sense of Hume's account of the artificial virtues. According to virtually everyone, Hume's discussion of the artificial virtues – and especially of the conventions on which he argues they depend -- is inspired, rich, and subtle. At the same time, also according to virtually everyone, Hume's discussion is deeply puzzling. Some have thought the puzzles so deep as to render Hume's position internally inconsistent or, if not, at least disingenuous.² Puzzling though Hume's discussion is, I hope to show that his account of the artificial virtues is not just consistent and sincere but plausible and attractive.

Background

Hume begins his discussion of morals, in the *Treatise of Human Nature*, arguing that our capacity to mark moral distinctions is not due to reason alone, but depends crucially on our ability to feel moral approbation and disapprobation. Just as we would not be able to distinguish things by their color, were we entirely color blind, so too, Hume maintains, we would not be able to distinguish actions, sentiments or characters morally, were we entirely disengaged affectively.³ “[V]irtue,” Hume writes, “is distinguished by the pleasure, and vice by the pain, that any action, sentiment or character gives us by the mere view and contemplation” (T 3.1.2.11/475).⁴ And, pressing the same idea, he claims that “To have the sense of virtue is nothing but to *feel* a satisfaction of a particular kind from the contemplations of a character.” (T 3.1.2.3/471).⁵ That “particular kind” of satisfaction, on Hume's account, is a distinctive feeling of moral approval, which one feels thanks to the workings of sympathy, when one sets aside considerations of self-interest.

¹ I am grateful to Lorraine Besser-Jones, Don Garrett, and Gerald Postema for very helpful discussions of Hume on the artificial virtues.

² Marcia Baron, David Gauthier, Rachel Cohon, Stephen Darwall, J.L. Mackie, Barry Stroud, all find inconsistent commitments in Hume's position. Baron argues that the problems run deep enough that we should see Hume as perpetuating the noble lie that justice, and the other artificial virtues, are virtues, when he thinks they are not. Along similar lines, Gauthier argues that Hume's account commits him to an error, which Gauthier suspects Hume realized but worked to keep in the shadows. Meanwhile, Cohon, Darwall, and Mackie each argue that Hume must have silently dropped, or restricted, one of the commitments that otherwise, together, cause problems. See Baron's "Hume's Noble Lie: An Account of His Artificial Virtues," *Canadian Journal of Philosophy*, Volume XII, Number 3 (September 1982), pp. 539-555; Gauthier's "Artificial Virtues and the Sensible Knave," *Hume Studies*, Volume XVIII, Number 2 (November, 1992), pp. 401-428; Cohon's *Hume's Morality: Feeling and Fabrication* (Oxford: Oxford University Press, 2008); Darwall's *The British Moralists and the Internal 'Ought'* (Cambridge: Cambridge University Press, 1995); Mackie's *Hume's Moral Theory* (London: Routledge, 1980); and Stroud's *Hume* (London: Routledge, 1977).

³ Of course, assuming that there are differences between what is virtuous and what is vicious, and what is red and what is green, beyond their moral and color qualities, a person sensitive to those differences might well be able to distinguish among them. But the distinctions thus drawn will not be in terms of the moral, or color, differences.

⁴ References to Hume's *A Treatise of Human Nature* (abbreviated "T"), by book, part, section, and paragraph are to the Clarendon Edition edited by David Fate Norton and Mary J. Norton (Oxford: Clarendon Press, 2000), with page references following "SBN" that are to the earlier edition edited by L. A. Selby-Bigge, revised by P. H. Nidditch (Oxford: Clarendon Press, 1978). References to Hume's *An Enquiry Concerning the Principles of Morals* (abbreviated by "EM"), by book, part, section, and paragraph, are to the Clarendon Edition edited by Tom Beauchamp (Oxford: Clarendon Press, 2000), with page references following "SBN" that are to the earlier edition edited by L. A. Selby-Bigge, revised by P. H. Nidditch (Oxford: Clarendon Press, 1975).

⁵ The particular kind of satisfaction or pleasure he has in mind is that of moral approval, which arises "only when a character is considered in general, without reference to our particular interest" (T 3.1.2.4/472).

Significantly, although Hume appeals generally to feelings of moral approval and disapproval, he refines his view of which moral approvals and disapprovals are relevant to moral distinctions and why. He does this in order to capture accurately the moral judgments we actually make while also explaining the differences between (i) our merely feeling approval of some trait and our judging it to be a virtue and (ii) our judging a trait to be a virtue and it actually being a virtue. In places Hume does write as if thinking that something is a virtue is simply a matter of feeling approval towards it, holding, for instance, that morality “is more properly felt than judg’d of” (T 3.1.2.1/470).⁶ Yet, as Hume recognizes, the plausibility of his theory depends on it respecting, and being able to explain, the fact that we distinguish between *feeling* approving of someone and *judging* her to be virtuous (just as we distinguish between something appearing blue and our judging it to be blue). And he takes great care in doing so. One of the important refinements involves distinguishing moral approval (which is felt “only when a character is considered in general, without reference to our particular interest” (T 3.1.2.4/472)) from other kinds of approval. Another is the identification of a privileged set of circumstances, distinctive of what Hume calls the “General Point of View,” in which feelings of moral approval set the standard for whether something counts as virtuous, and so a standard for moral judgment. Some trait is a virtue, on Hume’s view, if, but only if, it would garner moral approval from the General Point of View. With the standard in place we can distinguish between what we happen to approve of, given our actual situation, and what we would approve of, were we to take up the General Point of View. For our purposes, though, we can leave aside these refinements and the details of Hume’s account of moral judgment, in order to focus on what, according to Hume, secures the relevant moral approval.⁷

Virtuous Actions and Virtuous Motives

According to Hume, the relevant feelings of moral approval and disapproval are directed fundamentally not at actions (even though we do evaluate actions morally) but at the motives and durable traits of mind and character that give rise to action. “Tis evident,” he maintains, “that when we praise any actions, we regard only the motives that produced them, and consider the actions as signs or indications of certain principles in the mind and temper. The external performance has no merit. We must look within to find the moral quality” (T 3.2.1.2/477).⁸ This view is as old as Plato and Aristotle, who emphasized that the moral standing of one’s actions depends not merely on what one does but on why one does it. The same idea, of course, is at the heart of Kant’s moral philosophy, who holds that the moral worth of an actions depends exclusively on the agent’s intention.⁹ Just as there is a crucial difference between those who merely act like friends, and those

⁶ These are all places where Hume is emphasizing the essential role of feeling or sentiment in our capacity to make moral distinctions against those who think that reason alone would be sufficient.

⁷ In what follows, when I write, for instance, of something securing the “relevant approval” I have in mind distinctively moral approval felt when taking up the general point of view. An important feature of Hume’s account is that we can think about whether some motive, action, or character trait would secure that approval without ourselves actually feeling approval. See my “On Why Hume’s General Point of View Isn’t Ideal – and Shouldn’t Be” *Social Philosophy & Policy*, Volume 11, Number 1 (Winter 1994), pp. 202-228, for a discussion of Hume’s account of moral judgment, and my “Hume on Practical Morality and Inert Reason,” in *Oxford Studies in Metaethics*, edited by Russ Shafer-Landau (Oxford University Press, 2008), pp. 299-320, for a discussion of the important, but I think often misunderstood, connection Hume claims to hold between morality and motivation.

⁸ Although we often “fix our attention on actions... these actions are still considered as signs; and the ultimate object of our praise and approbation is the motive, that produced them” (T 3.2.1.2/477).

⁹ Kant begins the *Groundwork* claiming that the only thing that can even be conceived of as unconditionally good is a good will; everything else of value, he maintains, depends for its value on being compatible with such a will. Of course Kant has a distinctive view of what constitutes a good will, but the driving intuition is that various actions, whatever

who are truly one's friends, a difference that turns on why they act as they do, there is a crucial difference – a moral difference – between those who merely act in the way a generous, kind, or just, person would act and those who are truly generous, kind, or just, a difference that turns on why they act as they do, and one that affects the moral standing of the action performed.¹⁰ The “ultimate object of our praise and approbation,” Hume claims, “is the motive, that produced them” (T 3.2.1.2/477).

Similarly, Hume sees our evaluation of characters as tied to the motives with which we associate them. Thus, for instance, those who are genuinely benevolent are those of whom being motivated by a concern to help others is characteristic, and those who are genuinely cruel are those of whom being motivated by a desire to hurt others is characteristic. As a result, distinctions among the various virtues and vices are, as Hume sees things, ultimately distinctions among corresponding motives, thought of broadly as “principles in the mind and temper” (T 3.2.1.2/477), which include not simply preferences or desires, but also passions, dispositions, and the other aspects of personality that give rise to actions.

What marks a person as benevolent or cruel, just or selfish, courageous or manipulative, Hume is emphasizing, is not what they do but why they do it – the features of their personalities that lead them to act as they do. For this reason, if we blame a person for failing to perform some action and then discover “that the virtuous motive was still powerful over his breast, tho’ check’d in its operation by some circumstances unknown to us, we retract our blame, and have the same esteem for him, as if he had actually perform’d the action, which we require of him” (T 3.2.1.3/477-8).

Indeed, it is a central tenet of Hume’s account that a person’s actions are merely “external signs” of what matters to virtue.¹¹ For each kind of action that we might see as virtuous, and each character trait we might count as a virtue, Hume holds that there must be “in human nature” some corresponding motive (that is, preference, desire, disposition, or other principle of mind or temper) that would lead to that action, and that is characteristic of that virtue, which secures our moral approval from the General Point of View.

The Sense of Duty and the Regard to the Virtue of an Action

Once we have such a motive in sight, and an understanding of the sorts of actions to which it would give rise, we can identify that kind of action as, say, benevolent, just, or courageous, even in cases in which the corresponding motive is actually absent. And a person might perform that kind of action without the virtuous motive. Yet in acting without the virtuous motive that person will not be acting as a benevolent, just, or courageous person would – that is, not benevolently, justly, or courageously – even though she does what a benevolent, just, or courageous person would have done. And, in lacking a virtuous motive, the person’s action will not be virtuous, even if it is the *kind* of action a virtuous person would have performed (albeit with different motives).

their actual effects, count as valuable, if they are, only in light of why they were performed. Immanuel Kant, Groundwork for the Metaphysics of Morals, edited by Thomas E. Hill and translated by Arnulf Zweig (Oxford: Oxford University Press, 2002).

¹⁰ Hume is well aware that we might be pleased that someone has acted, even when the action is not virtuous, as when, for instance, the actions benefits us or those for whom we are concerned. But there is an important difference, according to him, between being pleased something was done and seeing it as virtuous. And his point is that we will see the action as virtuous only as we would approve of the motive behind it as virtuous.

¹¹ “If any *action* be either virtuous or vicious, ’tis only as a sign of some quality or character. It must depend upon durable principles of the mind, which extend over the whole conduct, and enter into the personal character” (T 3.3.1.4/575).

As Hume notes, someone might actually be moved to perform the kind of action by a “regard to the virtue” of the action, that is, from the recognition that it is the kind of action a virtuous motive would lead one to perform. With this in mind, Hume points out that

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 3.2.1.8/479)¹²

Hume here talks specifically of a sense of duty, which (according to Hume) is the sense that a failure to perform an action would show a blamable lack of a virtuous motive.¹³ Yet not every action that would result from a virtuous motive is such that absence of the motive would be blamable. So not every virtuous action is, on Hume’s view, a duty, although all actions that are a duty are virtuous. This means that a person might recognize of some motive that it is virtuous, though not a duty, and may still be moved to act as those with that motive would act by that recognition, though this would not be a case of acting from a sense of duty. Hume captures both possibilities by talking about “a regard to the virtue” of an action (T 3.2.1.6/478 and T 3.2.1.9/480) or “a sense of its morality,” (T 3.2.1.7/479) which involve seeing the action as one a person with a virtuous motive would perform (whether or not failing to have the motive is blamable). Although Hume in the passage above highlights a case in which a person notices his lack of a virtuous motive, it is worth registering that a person might actually have the relevant virtuous motive while also recognizing that it is a virtuous motive, thus finding himself with both the first virtuous motive and the motive provided by a regard to the virtue of the action.

Importantly, acting from the sense of duty or a regard to the virtue of the action may well secure the relevant approval and so itself count as a virtuous motive, at least when people are right about what their duty is.¹⁴ Taking advantage of this, “the public instructions of politicians, and the private education of parents, contribute to the giving us a sense of honour and duty” that provides a motive to act as virtue requires (T 3.2.6.11/533-4, see also T 3.2.2.25/500, T 3.2.5.12/523, T 3.2.8.7/545-6).¹⁵ In fact, Hume claims, “...every one, who has any regard to his character, or who intends to live on good terms with mankind, must fix an inviolable law to himself, never, by any temptation, to be induc’d to violate those principles, which are essential to a man of probity and honour” (T 3.2.2.27/501).

¹² While Hume recognizes the motivating force of the sense of duty, he also appreciates, and offers an explanation of, the fact that something being forbidden often makes it more attractive. This happens, he argues, when what is forbidden is something we have an independent motive to do; the opposition of the passions involved “commonly causes a new emotion in the spirits... This new emotion is easily converted into the predominant passion and increases its violence... Hence we naturally desire what is forbid, and take a pleasure in performing actions, merely because they are unlawful” (T 2.3.4.5/421).

¹³ Hume observes that “we blame a father for neglecting his child. Why? because it shews a want of natural affection, which is the duty of every parent” (T 3.2.1.5/478).

¹⁴ Presumably if a person’s sense of duty goes too far wrong, leading people to do what is actually contrary to duty, our approval of their acting on their sense of duty will find its limits.

¹⁵ At the same time, Hume rejects the idea that our sense of virtue is entirely the product of educators and politicians on the grounds that their efforts have an effect only because people already have an independent sense of virtue (T 3.2.2.25/500). He also argues that the range of virtues we recognize, which includes virtues that do not redound to the public interest, are not well explained by appeal to the efforts of educators and politicians alone (T 3.3.1.11/578-9).

The Need For A First Virtuous Motive

Although a motive of duty or a (motivating) regard to the morality of an action may sometimes itself be virtuous, Hume argues that it is necessarily a secondary motive¹⁶ -- one that relies, for its intelligibility, on there being in human nature (even if not in the person in question¹⁷) some *other* motive to perform the act in question. The motives of duty and more generally of regard to the morality of an action, according to Hume, always presuppose some *other* motive in light of which the action in question counts as one's duty or as virtuous. Indeed, Hume advances the following as an "undoubted maxim": "*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 3.2.1.7/479).¹⁸

Why accept that maxim? Hume offers a fairly simple argument: To see some action as one's duty, as virtuous, or as morally good, is to see it as a sign of a virtuous motive. To see a motive as virtuous is to think it would secure the relevant approval (i.e. approval from the General Point of View). While the regard for the virtue of an action, or a motive of duty, might well secure that approval (and be seen as securing it), to have such motives involves seeing some motive for an action as being such that it would secure the relevant approval. And, on pain of circularity, that must be some motive other than a regard for the action's virtue or the sense that so acting is one's duty: "To suppose, that the mere regard to the virtue of the action, may be the first motive, which produc'd the action, and render'd it virtuous, is to reason in a circle" (T 3.2.1.4/478). (This has come to be called the "circle argument.")

Hume does not say why the circularity would be a problem, taking for granted that it obviously is. But it is worth noting that the circularity poses two distinct problems.

First, if we were caught in a circle, the regard to the virtue of an action (or the motive of duty) would have no determinate content and so no particular actions that it would serve to motivate. Which actions are virtuous? Which are our duty? If, as Hume has argued, these questions are answered by appeal to the virtuous motive one might have in performing them, and that motive is simply a regard to the virtue of the action, we would have no way to identify which actions are to be so regarded and so no motive, yet, to do anything in particular. The problem is much like the one I would face if my sole desire were to fulfill the desires of my beloved, while her sole desire was to fulfill mine. With nothing more to go on, I would have no motive, yet, to do anything in particular. For me to have a motive to do something, my beloved needs to have some desire that does not refer exclusively back to my desire to satisfy her desires (or, at least, I have to think she has such a desire).¹⁹ What is needed is some desire of hers (that I might act to satisfy) other than a desire to satisfy my desire to satisfy hers. Similarly, when it comes to a regard for *the virtue of an action* (which involves seeing the

¹⁶ Considering actions prompted by humanity, Hume claims that the motive of "humanity bestows a merit on the actions. A regard to this merit is, therefore, a secondary consideration, and deriv'd from the antecedent principle of humanity, which is meritorious and laudable" (T 3.2.1.6/478)

¹⁷ Hume makes clear that while the motive needs to be one humans can and do have -- a motive in human nature -- a particular person might well lack the motive while still having a duty to act as the motive, if only he were to have it, would lead him to act.

¹⁸ Focusing on duty, he later makes the same point writing that "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" (T 3.2.5.6/518-19).

¹⁹ Alternatively, I might have some other desire with determinate content, that she might work to satisfy, with my help; then I would have a motive to do what would help her do what she can to satisfy that desire of mine. Either way, neither of us will have a motive to do anything in particular until there is some other desire in play.

action as one a person with a virtuous motive would perform). What is needed is some motive to perform an action, other than a regard to the virtue of that action – a motive that is virtuous (or, at least, such that I think it is virtuous). Once another motive is in play, in light of which I might see some particular action as virtuous, a regard for virtue will have a focus and direction, whereas without it, I would have neither.²⁰

Second, if we were caught in the circle we would find ourselves with no account of the virtuousness of the actions we are hoping to explain. The virtuousness of those actions, Hume maintains, depends on the virtuousness of the motive we have to perform them, but the motive on offer – the sense of duty (or a regard to the virtue of the action) – will be an approvable motive (that is, a virtuous motive) only if it is accurate -- that is, only if the action in question is really a duty or virtuous. But that means we need an independent explanation of the virtue of the action, an explanation that appeals to something other than a regard to the action's virtue. In order for a regard to the virtue of an action to be virtuous, "the action must be really virtuous; and this virtue must be deriv'd from some virtuous motive: And consequently the virtuous motive must be different from the regard to the virtue of the action." (T 3.2.1.4/478).²¹

Despite Hume's phrasing, the argument does not turn on holding that an "action must *be virtuous*, before we can have a regard to its virtue" (ital. added). If it did, then a fair response would be to deny that claim. After all, people might well regard some action as virtuous, perform it for that reason, and yet be wrong. Examples of misguided acts of duty are legion.

Yet this is to mistake how Hume's argument works. It does not assume that when we regard some action as virtuous we are always right. Hume can easily acknowledge that people may mistakenly think that an action is virtuous and perform the action out of a (wrongheaded) regard for its (supposed) virtuousness. What Hume is supposing, reasonably, is that a wrongheaded regard for the virtue of an action would not secure the relevant approval and so would not count as a virtuous motive.²² But suppose he is mistaken about that, and a wrongheaded regard for the virtuous of an action is virtuous. Even then, appealing to the mere regard for the virtue of an action will not work to explain the standing of a whole range of virtues as virtues.

On the one hand, that regard, precisely because it can be inaccurate, often leads people to act contrary to these virtues. The regard to virtue or the motive of duty, if not restricted to instances when the agent has things right, will lead to actions a person with the relevant virtue would never perform. So this regard cannot be the motive that explains the virtuousness of the actions that are distinctive of the virtues we are trying to explain. This is true even if a regard to the virtuousness of an action renders that action, whichever action it is, and however misguided the regard, virtuous. In that case, while there would be a virtue of, as we might put it, dutifulness, we would still have no explanation of the virtuousness of benevolence, prudence, justice, fidelity to promises, etc. as

²⁰ One way to avoid the force of this argument is to reject the idea that in order to regard an action as virtuous we need to see it as a sign of a virtuous motive. So, for instance, one might hold that, when it comes to the artificial virtues, conventionally established rules puts us in a position to identify the virtuous actions without having to appeal to any virtuous motives. More about this below, when we turn to the puzzling aspects of Hume's account of the artificial virtues.

²¹ The first is an epistemological difficulty, and would be solved if only there were a way to identify which actions are virtuous other than by appeal to the virtuous motive. The second is an explanatory or metaphysical difficulty.

²² It might also be reasonable to hold that if one acts as one sincerely believes to be required by duty one is acting virtuously. So, it is important that the force of Hume's argument does not depend on accepting what he assumes, as I go on to note.

contrasted with all the traits that are not virtues. At the same time, it is plausible to think, as Hume seems to, that being moved by a regard to the virtuousness of actions where that regard goes astray (especially if it goes far astray, to the point of counting deeply vicious actions as virtuous) is no virtue at all.

On the other hand, if we restrict ourselves to cases where the regard is accurate -- where the action really is a virtuous one -- we need an account of what makes the action virtuous, an account that cannot appeal to the (as yet not established as accurate) regard. And again the appeal to the regard for the virtuousness of the action leaves us without the explanation we need.

While Hume's argument invokes his account of what it is to see a motive as virtuous, the argument carries over to other views. Consider Kant, who famously maintains that we can and should act from the motive of duty, not from inclination. Significantly, Kant didn't hold that what matters is, as Hume would characterize it, a "mere regard to the virtue of an action." Instead, Kant sought a standard of duty that would give the thought that some action is one's duty content and direction. To that end, he argued that the standard of duty requires that our maxims meet certain formal and substantive requirements: that they be universalizable and such that in acting on them we treat humanity as an end. To regard an action as one's duty, according to Kant, is to see it as conforming to this standard. Yet this means that we are not talking about "a mere regard for duty," but rather a concern for whether one's maxims are universalizable and compatible with treating humanity, whether in oneself or others, as an end (and never solely as a means). This concern provides a way to avoid the circle Hume identified, and so the two problems that circle poses. At the same time, though, Hume would hold that it does so by offering a motive that goes beyond a (mere) regard for the virtue of the action in question: a regard to whether one's maxims are universalizable and such that acting on them involves treating humanity as an end. Whether such a concern is a virtuous motive, let alone (as Kant seems to hold) the only, or most fundamental, virtuous motive, is of course a substantive and controversial issue. What is important here, though, is that Kant's defense of the motive of duty involves identifying (what in Hume's terms would be) another motive, one that is not a mere regard to duty.²³

In any case, Hume's argument is that if actions count as virtuous thanks to the motives of which they are a sign, and the motives count as virtuous because they would secure moral approval, for each virtuous action there must be some corresponding motive (other than a sense of duty or a regard for its morality), of which the action is a sign, that secures our approval.²⁴ And the argument generalizes to those who reject the idea that motives count as virtuous because they secure the relevant approval. As long as one offers an account of what makes motives virtuous that does not reduce to holding that thinking a motive virtuous makes it so, a version of Hume's argument will

²³ The point is complicated by the fact that Kant treats thinking that an action meets this standard as one and the same with thinking, at least in the context of an imperfectly rational will, that it is one's duty. Likely Hume would argue both (i) that one might regard something as a duty without thinking it meets Kant's standard and (ii) that one might recognize, and be moved by the fact, that some action would meet the standard, without thereby thinking that so acting is one's duty. If Kant is right, Hume would argue, it is in being moved by such a thought that one does one's duty, not in doing what one believes is one's duty.

²⁴ One might think that Hume supposes that for each type of virtuous action that we recognize there is a distinctive virtuous motive that is not also the motive that underwrites the virtuousness of other kinds of actions. Yet Hume's views don't commit him to this. It might be that the very same motive -- say the motive provided by a concern for the welfare of others -- finds expression in a number of ways that we end up distinguishing as distinct virtues. There must be a virtuous motive -- that is, a motive of which we approve -- in order for the actions to count as virtuous, but (for all Hume argues) it might be that different kinds of actions get their standing as being virtuous from the same motive.

establish the need for some virtuous motive, other than the motive of duty or a regard to the virtue of the action, in order to account for the virtuousness of the action, and the virtuousness of the corresponding character trait (assuming that the virtuousness of actions and of character traits are a function of the virtuous motives they reflect).

Natural and Artificial Virtues

Hume holds that we can distinguish, among virtuous motives (i.e., motives that secure the relevant approval) between those that are artificial, and those that are natural. This difference turns on whether people can have the relevant motives absent the existence of various conventions and artifices that reflect the “designs, and projects, and views of men” (T 3.1.2.9/474); if they can, if the motives “have no dependence on the artifice and contrivance of men” (T 3.3.1.1/574), then they are natural, otherwise the motives are artificial.²⁵ This difference among the motives underwrites a distinction among corresponding virtues between those that are natural and those that are artificial.

All of the virtues, Hume argues, fall into (at least) one of four categories: they are all either useful, or immediately agreeable, to the person who has the virtue, or to others. And he argues that they all secure the relevant moral approval because, thanks to sympathy, our idea of the pleasure the virtues produce (by being either useful or immediately agreeable) are transformed into a pleasure, that in turn, gives rise to approval. “When any quality, or character, has a tendency to the good of mankind, we are pleas’d with it, and approve of it; because it presents a lively idea of pleasure; which idea affects us by sympathy, and is itself a kind of pleasure” (T 3.3.1.14/580). Hume offers a detailed account of how this works. But we can leave aside the detailed story. What is important here is the difference, among the qualities or characters that (via sympathy) give rise to the relevant moral approval, between those that are *natural*, and those that are *artificial*.

The natural virtues Hume has in mind include “beneficence, charity, generosity, clemency, moderation, equity” (all of which count as social virtues because of “their tendency to the good of society” (T 3.3.1.11/578-9)) as well as “prudence, temperance, frugality, industry, assiduity, enterprize, dexterity”²⁶ (which make those who have them “serviceable to themselves, and enable them to promote their own interest” (T 3.3.1.24/587-8)) and also good humour, wit, eloquence (which are immediately agreeable either to the person himself or to others (T 3.3.4.8/611)). What makes these virtues natural is that the motives characteristic of people with them are available, and secure the relevant approval, independently of what particular artifices and contrivances, if any, might be in place. (Of course, how a charitable person acts will depend, among other things, on the various conventions she is surrounded by. But her virtuous motive – her concern to help those in need -- is available regardless of those conventions. The same is true, Hume holds, of all the other natural virtues and their corresponding motives.)

²⁵ Mackie suggests that the artificial virtues are also such that “it is only by some artifice or invention that people have come to feel approval” of the motives characteristic of those with the artificial virtues. *Hume’s Moral Theory*, p. 77. Yet Hume explicitly denies this, holding that after the motives become available (thanks to the artifice and invention of man) “it is *naturally* attended with a strong sentiment of morals” (T 3.3.1.12/579-80). At the same time, Hume clearly thinks that politicians and educators work to strengthen the sentiments. Speaking of the rules of justice, Hume notes that while, “the sense of morality in the observance of these rules follows *naturally*, and of itself, tho’ ’tis certain, that it is also augmented by a new *artifice*, and that the public instructions of politicians, and the private education of parents, contribute to the giving us a sense of honour and duty in the strict regulation of our actions with regard to the properties of others” (T 3.2.6.11/533-4).

²⁶ As well as “...*perseverance, patience, activity, vigilance, application, constancy*” and “other virtues of that kind, which ‘twill be easy to recollect...” (T 3.3.4.7/610-11).

In contrast, Hume maintains, there are a number of virtues that depend on motives that we can have only thanks to human “artifice and contrivance” and, in particular, only in the context of *voluntary* conventions. He has in mind, as paradigms, justice concerning property, the performance of promises, and allegiance to government, although he also discusses, more briefly, the duties of princes and political ministers, chastity, and modesty. Despite their evident differences, Hume considers them all to be matters of justice. (The title of the section in which all the artificial virtues are discussed is “*Of Justice and Injustice*.”)

Exactly what leads Hume to this broad conception of justice is unclear. But the conception in play seems to be one of justice as requiring restraint in the face of what, absent conventions and the motives they make possible, would be temptation, where the failure of restraint is blamable. Noting that “the avidity and partiality of men wou’d quickly bring disorder into the world if not restrain’d by some general and inflexible principles” Hume argues that it was “with a view to this inconvenience, that men have establish’d those principles, and have agreed to restrain themselves by general rules, which are unchangeable by spite and favour, and by particular views of private or public interest” (T 3.2.6.9/531-3).²⁷ This general story fits his account not simply of justice as it relates to property, but also his accounts of the other artificial virtues. In each case, Hume argues, the virtues consist in a regard for principles or general rules that “men have establish’d,” through convention, as restraints.²⁸ What makes all these virtues artificial is that the motives characteristic of people with them are available, and secure the relevant approval, only in contexts shaped by particular artifices, inventions and conventions.²⁹

It is worth emphasizing that both the motives, and the approval they garner from the General Point of View, depend on conventions. The motives are not even available, Hume argues, absent established conventions – there is nothing in particular one is required to do, for instance, concerning property, until there are rules in place distinguishing property that marks the difference between mine and thine. But even when they are in place, Hume argues, their “tendency to the good of mankind” depends on “the whole plan or scheme,” that is, on the convention, of which they are a part (T 3.2.2.22/497-8). The benefits they secure are a product of collective action, not available absent the collaboration of others. In contrast, the natural virtues are such that both the motives characteristic of those with them motives, and the approval these motives garner from the General Point of view, do not depend on conventions.

²⁷ While the rules are “unchangeable by spite or favour, and by particular views of private or public interest” they are otherwise changeable as circumstances and mutual interest might recommend. This sensitivity to circumstance, including history, and mutual interest, explains why the specific actions that might count as virtuous can differ dramatically through time and across communities.

²⁸ One might think that the failure of restraint is blamable because others have a right to the restraint. This fits well with Hume’s discussion of property, promises, allegiance, and even the duty of princes and political ministers, where it is easy to identify those who have a right to restraint. But it fits less well with his account of the virtues of modesty and chastity, which do not seem tied to someone having any particular right to expect the restraint. While restraint is key, it seems that the restraint is not something to which others always have a right, even if a failure to exercise it is, in Hume’s view, blamable.

²⁹ Needless to say, it is possible to imagine situations in which the motives characteristic of, say, charitable people will prove to be neither useful nor agreeable to those who are charitable nor to others. Were the world such that this was generally true, Hume would hold (most likely) that, in such a world, charity is not virtuous or (less likely, but still consistently) that such a world would be one in which virtue turns out not to be either useful or agreeable. What is important to Hume’s account of morality is that, as the world is, the virtues are useful or agreeable, either to the possessor or to others. It is this feature of them that allows him to explain the moral distinctions we draw (between, say, virtue and vice) by appeal to our moral sentiments of approval and disapproval.

Although the artificial virtues depend, for their possibility, on human conventions, Hume emphasizes that those conventions are nearly inevitable. “Mankind is an inventive species; and where an invention is obvious and absolutely necessary, it may as properly be said to be natural as any thing that proceeds immediately from original principles, without the intervention of thought or reflexion. Tho’ the rules of justice be artificial, they are not *arbitrary*” (T 3.2.1.19/484).

As becomes clear, a crucial feature of all the artificial virtues is that they each solve salient shared problems that we would otherwise face.³⁰ The differences among the particular artificial virtues reflect differences among the problems and so differences in how they might be solved. Thus, for instance, justice concerning property addresses the need we have for stable possessions and an effective way to transfer goods, while fidelity to promises enables us to undertake commitments to future performance that others can trust, and the allegiance to political authority makes stable government possible in a way that it otherwise would not be.³¹

To make the case that justice, understood as respect for the property of others, is an artificial virtue, Hume first considers the natural motives one might think of as candidate motives to justice: private interest, public benevolence (or regard to “publick interest”), and private benevolence. Each of these motives does often prompt people to perform just acts (that is, the kind of acts a just person would perform). However, Hume points out, they are each such that under some circumstances people acting on these motives would act contrary to the requirements of justice. Sometimes acting *unjustly* is in one’s self-interest, sometimes it will promote the general welfare, and sometimes it will be advantageous for those about whom one cares. Moreover, it is worth noting, even when these motives lead one to act as justice requires, a person so motivated is not one whom we admire as just, but rather (if we admire the person at all) as, prudent, public spirited, or kind, respectively. Hume does not emphasize this second point. But a crucial requirement on any motive that might be the first virtuous motive to justice is that it is the motive we admire as characteristic of those who are just.³² In light of this, and assuming he has canvassed and found wanting all the plausible natural motives to justice, Hume concludes “that we have naturally no real or universal motive for observing the laws of equity but the very equity and merit of the observance” (T 3.2.1.17/483).³³ What is needed is some motive for just actions -- other than a regard for their morality -- that is both such that it would consistently lead people to act in ways we recognize as just and is the motive we

³⁰ This is true, as well, of the natural virtues (at least those that are useful either to the possessor or to others). But how they solve the relevant problems is importantly different. The natural virtues work to solve the problems unilaterally, so to speak, independently of what others might do, while the artificial virtues work to solve the problems they do only thanks to the collaboration of others. See “Hume and the Bauhaus Theory of Ethics,” for a discussion of connection, on Hume’s view, between salient practical problems and the virtues. *Midwest Studies in Philosophy*, Vol. XX (University of Notre Dame Press, 1996), pp. 280-298.

³¹ And the virtue of chastity, on Hume’s telling, encourages confidence in men that certain off-spring are their own, thus inducing them to provide support and protection both for those children and for their mothers.

³² Rachel Cohon rightly emphasizes this point in her “Hume’s Difficulty with the Virtue of Justice,” *Hume Studies*, Volume 23, Number 1 (1997): 91-112, pp. 104-105.

³³ This passage continues by rejecting a regard for the equity and merit of observance as a motive: “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.” The “naturally” in this conclusion did not appear in the first edition of Book III of the *Treatise*, although Hume added it in his own copy. See P.H. Nidditch’s “Textual Notes” to his revised edition of Selby-Bigge’s edition, p. 669. There is a parallel passage, regarding the motive to the performance of promises (“’tis evident we have no motive leading us to the performance of promises, distinct from a sense of duty” (T 3.2.5.6/518-19) in which “naturally” was not indicated by Hume to be missing. Yet a few sentences later “naturally” shows up: “as there is naturally no inclination to observe promises, distinct from a sense of their obligation” (T 3.2.5.6/518-19). It’s clear, I think, that “naturally” belongs in all three places.

admire as the motive of a just person.³⁴ And that motive, Hume has argued, must be non-natural, that is *artificial*, motive since no natural motive fits the bill.

Similar considerations would establish too that the keeping of one's promises is an artificial virtue. After all, it is clear that the motives Hume considered in discussing justice would likewise fail consistently to lead one to keep one's promises. And it is clear too that these are not the motives we approve of when we admire someone for keeping her promise (even if, when these motives lead someone to keep her promise, we are grateful she did). When these motives lead someone to keep her promises, we would not see her actions as a reflection of the virtue of someone who "true to her word" (T 3.2.5.11/522-3). It is one thing to keep one's promise because it happens to be in one's interest, quite another to do it because one gave one's word.

When it comes to promises, though, Hume focuses his attention on arguing "*that a promise would not be intelligible, before human conventions had establish'd it; and that even if it were intelligible, it won'd not be attended with any moral obligation*" (T 3.2.5.1/516). In defense of the first claim, Hume maintains that "If promise be natural and intelligible, there must be some act of the mind attending these word, *I promise*; and on this act of the mind must the obligation depend" (T 3.2.5.2/516).³⁵ He is thinking of the act of the mind, whatever it might be, as what would provide the necessary "*inclination to perform [i.e. motive], distinct from a sense of duty*" (T 3.2.5.7/519) that would in turn secure the relevant moral approval. Hume then argues that whatever is involved in making a promise, no particular act of the mind is required. Making a promise is not a matter of resolving, Hume points out, nor of desiring, nor of willing, to act in a certain way. On the one hand, a person might perform any of these acts and yet manifestly not have made a promise; on the other hand, a person might well have made a promise without actually resolving, desiring, or willing, to perform as promised. These acts of the mind are neither necessary nor sufficient for having made a promise.

What candidates are left? Perhaps, Hume suggests, one might think that it is "the *willing* of that *obligation*, which arises from the promise" (T 3.2.5.3/516). Against this possibility, he argues that a mere act of will can change neither our own, nor anyone else's, sentiments of approval. Yet a "change of the obligation supposes a change in sentiment; and a creation of a new obligation supposes some new sentiment" (T 3.2.5.4/517). As a result, he holds, it would be absurd "to will any new obligation...nor is it possible," he thinks, "that men cou'd naturally fall into so gross an absurdity" (T 3.2.5.4/517). In putting his argument in this way, Hume suggests that while men couldn't *naturally* fall into so gross an absurdity, they might, thanks to the "artifices and contrivances of men," fall into such an absurdity. And, in fact, it seems clear that Hume believes that we can and do fall into the absurdity of thinking we can will an obligation. Yet, precisely because willing an obligation is an absurdity, it cannot be the act of mind on which a promise actually depends. Nor, even supposing such an act of the mind, could it actually establish an obligation, for the same

³⁴ It is worth noting that a regard to equity, as Hume is setting things out here, is not a moral motive, although a regard to the *merit* of observing equity is. A regard to equity is a recognition that certain actions are called for, or required by, certain principles or general rules. Having discussed the origin of such principles and rules, Hume turns separately, and later, to the question of "*Why we annex the idea of virtue to justice*" (T 3.2.2.23/498), making clear that he thinks we might recognize something as just, or a matter of equity, without thereby seeing it as virtuous. The idea of virtue is annexed to the idea of justice, not contained in it, though justice is indeed, Hume thinks, a virtue.

³⁵ Importantly, Hume's point is not that the meaning of "I promise" is fixed by conventions (although it's meaning is fixed by conventions). His point is that the virtuous motive a person acts on, if she has the virtue of being true to her word, itself depends on there being certain conventions. The alternative he considers, and argues against, is that there is some state of mind, available independently of conventions, that constitutes the making of a promise and that gives rise to a motive to keep the promises made.

reason: A new obligation would require new sentiments, but an act of willing, by itself, creates no new sentiments.³⁶

Again assuming he has canvassed, and found wanting, all the plausible natural motives to keep one's promise, Hume concludes: "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" (T 3.2.5.6/518-19).³⁷ What is needed is some motive to keep promises -- other than a regard for their morality -- that would consistently lead people to keep their promises while also being the motive we admire as the motive of a person who "keeps her word." And that motive, Hume has argued, must be a non-natural, that is *artificial*, motive since no natural motive fits the bill.

Although Hume spends much less time discussing the other artificial virtues -- allegiance to government, the morality of princes and political ministers, chastity, and modesty -- considerations of the same sort apply. In each case, as with justice and fidelity to promises, Hume holds that there is no *natural* motive (other than a regard to their virtue) that would consistently lead people to perform the actions we recognize as virtuous. Moreover, the natural motives that do sometimes prompt such actions are not the motives we admire as characteristic of those with the relevant virtues. For these virtues, as for justice and fidelity to promises, the corresponding virtuous actions require "real and universal," but *artificial*, motives.

If it is virtuous to act justly, to keep one's promises, to give allegiance to one's government, etc., and if each virtuous action requires a virtuous motive, there must be some virtuous motive, other than regard to the virtue of so acting, to perform such actions. Yet there is no natural motive that both consistently leads people to act as the virtue requires and that is admired as characteristic of those with the virtue in question. Since the motive of duty (or a regard to the morality of the actions), won't do (in light of "the circle argument"), and there is no suitable natural motive, then there must be other -- artificial -- motives, due to "the artifice and contrivance of men," that serve as the first virtuous motives for the corresponding virtues.

Which Artifices?

Hume's discussion of which artifices matter to the artificial virtues is both lovely and seminal.³⁸ In

³⁶ "A new obligation supposes new sentiments to arise. The will never creates new sentiments. There could not naturally, therefore, arise any obligation from a promise, even supposing the mind could fall into the absurdity of willing that obligation" (T 3.2.5.5/518).

³⁷ Here "naturally" does appear in the first edition, lending support to the thought that it belongs in the passage concerning justice that Hume amended in his copy, and belongs as well in a passage concerning promises, a page earlier, where Hume writes "'tis evident we have [naturally] no motive leading us to the performance of promises distinct from a sense of duty" (T 3.2.5.6/518-19), although it doesn't appear there and was not added later by Hume. This matters, as will become clear, to whether Hume is arguing that there is *no motive whatsoever* to act justly, or to keep one's promise, other than the motive of duty, or that there are *no natural motives* to perform such actions. For now, I just note that, as I interpret Hume's position, the restriction to natural motives belongs in and is important.

³⁸ Modern game theory has its roots in Hume's discussion of the "origin of justice and property" and more generally in his account of the origin of the conventions associated with the various artificial virtues. Although Hume sometimes writes as if the problems that give rise to the conventions are simply matters of coordination, especially as he emphasizes that all can expect to benefit from having some convention or other in place, he is clearly sensitive both to the differential impact of different conventions on peoples' interests, and on the incentives people will have to fail to comply (perhaps secretly) when others are complying. Concerning the latter, he recognizes various forms of enforcement as crucial to the stability of the conventions, even as he emphasizes that, compared to life without the conventions,

it, Hume identifies the relevant conventions as depending on “a general sense of common interest; which sense all the members of the society express to one another, and which induces them to regulate their conduct by certain rules” (T 3.2.2.10/490). As Hume turns his attention from one artificial virtue to the next he identifies the (different) interests we have in common that are served by having various conventions; some to define property, some to make promising possible, some to establish governments to which allegiance might be due. In each case, Hume argues, we are all better off with the conventions than we would be without and we all have an interest in conforming to the conventions, provided others do so as well. Precisely which conventions we happen to put in place matters much less, he thinks, than that we have some conventions rather than none. And we can have the conventions at all only if people succeed in regulating their conduct by common rules or principles.

Thinking specifically of the rules of justice concerning property, Hume offers this description:

I observe, that it will be for my interest to leave another in the possession of his goods, *provided* he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually expressed, and is known to both, it produces a suitable resolution and behaviour. And this may properly enough be called a convention or agreement betwixt us... since the actions of each of us have a reference to those of the other, and are performed upon the supposition, that something is to be performed on the other part (T 3.2.2.10/490).

When it comes to property, the common interest is based, he argues, on the fact that we find ourselves in a situation in which there is moderate scarcity (of the things we need and want) as well as limited generosity among our fellows. This combination gives point to finding a way to make possessions secure, which an effective system of property will do. If either we enjoyed a superabundance of goods, or a community of people who were completely unselfish, he argues, the need for property would disappear. But we are not so lucky and justice finds its place.

Parallel accounts hold for the other artificial virtues, each finding its place thanks to our being in circumstances in which establishing a convention offers the possibility of mutual advantage. In the case of promises, the circumstances are those in which we would benefit not only by being able to express a willingness, or an intention, to do something in the future, but by being able to make a commitment to future performance. This requires, Hume argues, a convention that enables us to undertake a commitment (e.g. give our word) in a context that makes us liable to a penalty of some sort for a failure to perform. In the case of allegiance, the circumstances – different from those required for property and promises, but (Hume thinks) largely motivated by problems they render salient – are those that make establishing a government advantageous. In particular, and most importantly, the circumstances that matter are those we find ourselves in when we notice that while the conventions of property and promises are advantageous, we are liable to temptation to act unjustly and to break our promises. This liability, Hume thinks, is due primarily to the “natural infirmity” we all suffer of giving way “to the solicitations of our passions, which always plead in favour of whatever is near and contiguous” (T 3.2.7.2/534-5). The effect of this is that we are seduced by the prospect for short-term benefits at the cost of long-term advantages, not least the

everyone can expect to benefit by compliance, as long as others are complying as well. David Lewis’s *Convention* (Harvard University Press, 1969) is a careful articulation, development, and defense of Hume’s insights.

long-term advantages that come from successfully sustaining conventions that are mutually beneficial. Concerning duties of allegiance, Hume maintains that they rest “exactly on the same footing, and have the same source both for their *first invention* and *moral obligation*. They are contriv’d to remedy like inconveniences, and acquire their moral sanction in the same manner, from their remedying those inconveniences” (T 3.2.8.4/542-3).³⁹

At the heart of the conventions that make the artificial virtues possible is their being mutually advantageous and recognized as being so. Also crucial, though, is that the conventions, and so their advantages, are available only when others will act accordingly, as long as one does one’s part. Both the mutual advantage and the reciprocal performance are crucial elements to these conventions working to underwrite moral *virtues*.

Hume is of course well aware that not all conventions are mutually advantageous. Fashion comes to mind, as does (more significantly) slavery. Both take the shapes they do, and lead to the actions they do, only thanks to the “the artifice and contrivance of men.” Yet these artifices and contrivances work their effects in ways that do not depend on *mutual* advantage, even as they promote the advantage of some. And this, for Hume, disqualifies them from making a virtue of conformity. A similar point holds for societies in which there is a permanent underclass, forced to comply with conventions from which they do not benefit. On Hume’s account, those in the underclass have no duties of justice with regard to the established property of others.⁴⁰ For Hume, the conventions that underwrite artificial virtues are restricted to those that are mutually advantageous, where the advantages in question depend on people coordinating their behavior (i.e., acting on the “supposition, that something is to be performed on the other part.”)⁴¹ Significantly, on Hume’s view, it matters that the conventions in question actually be in place. That a convention *would* be advantageous, if only we had it, may be reason to work to establish it, but it is no reason to act according to it, absent the participation of others. While there are any number of different conventions of, say, property, that *would* be advantageous, the actual benefits of conforming to one or another usually depend on which happens to have been established. Of course, acting in accordance with rules that *would* be part of an advantageous convention may be a way of helping to establish the convention. If so, then one would have some reason to do so. But in acting so as to

³⁹ Here Hume is arguing that our duty of allegiance to government is not founded on our having made a promise, on the grounds that our obligations of fidelity to promises and of allegiance to government are “built on the very same foundation.” He allows that we can establish a convention that makes promising possible, and fidelity to promises obligatory, absent a government. That means we might actually promise allegiance. But whether we do or not, the account of our obligation to keep our promises carries over directly, without the interposition of a promise, to our obligation of allegiance to established governments (if they are suitably mutually advantageous).

⁴⁰ While it is clear Hume thinks that those who do not benefit from the convention have no duty of justice, it is unclear whether he thinks the convention establishes duties of justice among those who have benefitted, as long as their benefit comes from the willing compliance of others on the expectation of mutual compliance. It may be that he thinks that those who enjoy the benefits have duties to each other, in light of the mutual, and mutually dependent, advantages they’ve secured. Whether or not they do, they would have responsibilities – presumably weightier responsibilities – to those who have been exploited by the conventions in place. Yet it seems that on Hume’s view these responsibilities would not be matters of justice, at least if those who have been exploited can neither make their resentment felt nor enter into alternative arrangements that would benefit others. Consider Hume’s example of “a species of creatures intermingled with men, which, though rational, were possessed of such inferior strength, both of body and mind, that they were incapable of all resistance, and could never, upon the highest provocation, make us feel the effects of their resentment.” In that case, he concludes, “that we should be bound by the laws of humanity to give gentle usage to these creatures, but should not, properly speaking, lie under any restraint of justice with regard to them” (EM 3.18/190-1).

⁴¹ He pretty clearly holds that wherever actual conventions have provided mutual advantage, even if not equal advantage, duties of justice are found.

bring about the mutually advantageous convention one would so far be acting from benevolence or self-interest, not (yet) from justice (since, absent the convention, so acting will not count as an instance of justice).

In the telling, Hume often writes as if each of the relevant conventions are established self-consciously, and voluntarily, by all who are subject to them. And he clearly thinks that those who found themselves without conventions defining property, making promises possible, and establishing governments, would, realizing their advantages, put them in place and would then voluntarily conform to them -- assuming others did as well. The obvious value of such conventions, properly designed (as to be mutually advantageous), makes them something that people would all willingly establish and work to support.

Yet Hume is very sensitive not only to the fact that many conventions are not mutually advantageous, but also to the fact that many that are advantageous emerge slowly, often without conscious planning, even as their perpetuation depends on an (often nebulous) appreciation of their value.⁴² And they develop in ways that leave choices among different conventions as something we rarely face without finding ourselves already within conventions from which we have benefitted from the willing conformity of others. Once we find ourselves within mutually beneficial conventions, the range of available options is constrained by our established obligations and restricted to options that might be available without losing the benefits already in place. What matters most, Hume emphasizes, is that there actually be a mutually advantageous convention. After all, without one all sorts of significant advantages will be lost to all.

At the same time, Hume largely leaves to one side the significant complications that come with recognizing that different conventions would bring different advantages to different people, a fact that might be expected to introduce more than a little friction among different factions. Such problems would be especially pressing if the parties to the convention are supposed to settle on a convention together through deliberation or bargaining. That our actual conventions are largely ones we find ourselves in, and benefitting from, rather than ones that we establish *ex nihilo*, plays an important role in explaining why so much is taken for granted in practice.

Of course, sometimes clearly better conventions are so close to hand that an established convention might begin to lose its grip. Under such circumstances, I suspect, Hume's view would be that we should work to establish the better convention, but that until we succeed we have a duty to act according to the less than optimal arrangement that is in place, as long as it is mutually advantageous and has been sustained thanks to others doing their part on the expectation that others would as well.⁴³

⁴² When it comes to property, this easy establishment is due in large part, Hume thinks, to there being an obvious rule for settling, in the first instance, what belongs to whom: "it must immediately occur, as the most natural expedient, that every one continue to enjoy what he is at present master of, and that property or constant possession be conjoin'd to the immediate possession." Strikingly, Hume appeals to what has come to be called "the endowment effect" to explain the comfort we would feel with this solution: "Such is the effect of custom, that it not only reconciles us to anything we have long enjoy'd, but even gives us an affection for it, and makes us prefer it to other objects, which may be more valuable, but are less known to us" (T 3.2.3.4/503-4).

⁴³ The case of revolutions is nicely complicated. Hume thinks that there will likely be no rule, within established governments, giving permission to revolt. That means, on his account, that we will not have a right to revolt. But a government can lose its claim to allegiance (T 3.2.9.4/552-3). And if the government is bad enough, and if we can

That actions are performed as part of an established, and actually mutually beneficial, scheme, is essential to their counting as virtuous. At the same time, however, Hume emphasizes that the particular virtuous actions need not themselves be useful. “Taking any single act,” Hume notes, “my justice may be pernicious in every respect...” (T 3.2.2.22/497-8). In particular, he recognizes, it may be disadvantageous both to the person who performs it and to others.

When a man of merit, of a beneficent disposition, restores a great fortune to a miser, or a seditious bigot, he has acted justly and laudably, but the public is a real sufferer. Nor is every single act of justice, consider'd apart, more conducive to private interest, than to public (T 3.2.2.22/497-8)

What renders justice a virtue is that “however single acts of justice may be contrary, either to public or private interest, ‘tis certain, that the whole plan or scheme is highly conducive, or indeed absolutely requisite, both to the support of society, and the well-being of every individual (T 3.2.2.22/497-8). It is with an eye to the utility of the scheme, of which the act is seen as an integral part, that acts of justice garners the moral approval that constitutes them as virtuous.⁴⁴ Consider them apart from the mutually advantageous scheme, or consider acts that are a part of a scheme that is not mutually advantageous, and the grounds of approval disappear.

Stressing this idea, Hume contrasts the natural virtues with the artificial virtues claiming “that the good, which results from the former, arises from every single act, and is the object of some natural passion: Whereas a single act of justice, considered in itself, may often be contrary to the public good; and it is only the concurrence of mankind, in a general scheme or system of action, which is advantageous” (T 3.3.1.12/579-80). But this is, at least, misleading. Hume himself emphasizes of the natural virtues, no less than the artificial virtues, that they do not always produce the good at which they aim. In particular, for instance, sincere acts of benevolence do not always succeed in helping others, though they are virtuous. While, on Hume’s account, benevolence secures moral approval thanks to its usual effects, we approve of that motive even when it fails to have those effects. As Hume emphasizes, “virtue in rags is still virtue; and the love, which it procures, attends a man into a dungeon or desert, where the virtue can no longer be exerted in action, and is lost to all the world” (T 3.3.1.19/584). Hume’s point here, in contrasting the natural with the artificial virtues, is that the advantages, in light of which the artificial virtues secure approval, emerge only thanks to “the concurrence of mankind, in a general scheme or system of action” – a convention – whereas the advantages, in light of which the natural virtues secure approval, are available without recourse to conventions.

succeed in replacing it with something enough better to justify the significant hardships revolution imposes on a people, then we may be right to revolt even if we do not have a right to revolt.

⁴⁴ Of course not every act of justice is actually needed to sustain the advantageous scheme. What is at stake here is how we conceive of the act, when we take up the General Point of View. Just as, from the General Point of View, we think of benevolent acts as *a kind of act that is beneficial*, and so approve of particular instances of that kind regardless of whether they are actually useful, so too we think of just acts as *a kind of act that is an integral part of a scheme that is (mutually) advantageous*, and so approve of particular instances of that kind regardless of whether they are actually (mutually) advantageous. For a discussion of why, from the General Point of View, the focus is the usual effects of what is being judged, that is on the effects of the kind in question, rather than the actual effects of the specific instances, see my “Why Hume’s General Point of View Isn’t Ideal – and Shouldn’t Be,” *op. cit.*.

The Missing Motive

A striking and puzzling fact about Hume's discussion of the artificial virtues is that he never explicitly identifies the first non-moral yet virtuous motive(s) to act as the artificial virtues require. This is despite having taken a lot of time to argue against various candidate motives, while also arguing that for every virtue there must be some such motive – that is a motive, other than the a motive of duty (or a motive depending on the recognition that so acting is virtuous). Hume doesn't, of course, doubt that we can be motivated by a sense of duty. But he does think that the motive of duty, if it is not misguided, depends on there being some other, non-moral, motive, which itself secures the approval that constitutes it as virtuous. Hume never clearly and explicitly tells us what that motive is.⁴⁵

This has led people to offer a range of interpretations. According to one

There is no first, non-moral yet virtuous, motive for the artificial virtues.

On this line view, Hume is silent about the first – non-moral yet virtuous – motive(s) for the artificial virtues because he thought no such motive exists. Recommending this idea is not just the fact that Hume seems never to explicitly identify such a motive, but that he works so carefully to rule out what look to be all the possible candidates.

If this is right, then there are two paths to follow. The first involves holding that

The artificial virtues are not genuine virtues.

This involves holding firm to Hume's claim that for each virtue there must be a corresponding non-moral yet virtuous, motive, that renders the corresponding acts virtuous, and concludes that Hume's view must have been that the artificial "virtues" are *not actually virtues at all*. They may, in various ways, resemble virtues, not least in being useful to society, but they are, nonetheless, not genuine virtues, on this interpretation, precisely because there is no first virtuous motive to their performance. No doubt people are (sometimes) motivated to act as the "virtues" would require, but this is always either because of one of the motives that Hume has rightly rejected as candidate motives for the "virtue" in question, or because they have been convinced by politicians, educators, and parents that so acting is a duty.⁴⁶ Still, that people act justly regularly benefits society. So it is no wonder that politicians and others encourage the (false, on this interpretation) idea that so acting is virtuous. In fact, the benefits that come from convincing people that justice is a virtue might help explain why Hume was reticent to emphasize that there was no non-moral yet virtuous motive.

One might have expected Hume to make this point, if it were his view. He doesn't. Yet, as defenders of this interpretation point out, the subversive nature of the thesis – that justice is not (actually) a virtue – may have led Hume to down play the point in favor of joining the forces of those who are promulgating the fiction that it is a virtue.⁴⁷ Hume does after all argue that "though the philosophical truth of any proposition by no means depends on its tendency to promote the

⁴⁵ For a detailed discussion of this puzzle see Don Garrett's "The First Motive to Justice; Hume's Circle Argument Squared," *Hume Studies*, Vol. 33, Number 2 (2007), pp. 257-288.

⁴⁶ See T 3.2.2.25/500.

⁴⁷ Marcia Baron defends this interpretation in "Hume's Noble Lie: An Account of His Artificial Virtues," *Canadian Journal of Philosophy*, Vol. XII, Number 3, September 1982, pp. 539-555.

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” (EM 9.14/278-9). He may well have been concerned that the truths he uncovered would be dangerous and pernicious, were he to make them explicit.

Yet this interpretation strikes me as not especially plausible. To think that kindness, benevolence, and the other natural virtues are genuine virtues, and are virtues because they are useful or agreeable to self or others, while holding that justice (as it relates to property, to keeping one’s word, to giving allegiance to one’s country), which is useful to others, would fail to be a virtue, seems just the wrong way to go. If Hume really thinks there is no first, non-moral yet virtuous, motive to justice, then the most plausible move is to reject Hume’s claim that every virtue is such that there is such a corresponding motive.

This reaction recommends the second path one might follow, if one believes Hume thought there was no first, non-moral yet virtuous, motive for the artificial virtues. On this interpretation

Not all virtues require a first, non-moral yet virtuous, motive.

Perhaps the mark of the artificial virtues, as opposed to the natural virtues, is that there is no such motive, and the mistake is to think that what is true of an important class of virtues – the natural virtues – is true of all. The suggestion involves holding that, when it comes to the artificial virtues, Hume abandons, or at least should have abandoned, the “undoubted maxim” mentioned earlier: “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 3.2.1.7/479).

What is striking here is that Hume seems never to have rejected this maxim. Nor is it clear what in his argument for it he could consistently reject, given his over all account of the nature of virtue. Embracing this second interpretation requires not simply restricting his maxim to the natural virtues, but explaining how his general account of virtue survives the restriction. There are two elements of his general account that might be jettisoned: either (i) Hume’s commitment to the virtue of an action depending on the motive with which it is performed or (ii) his commitment to holding that the motive of duty cannot be the first (albeit moral) motive, the approval of which constitutes an action as virtuous. At least one would have to go.

The first commitment is one that Hume never argues for, yet it is deeply intuitive and has a long and respectable pedigree going back to the ancients. The second commitment is more distinctive of Hume, but it rests on an explicit argument – the circle argument -- for thinking that neither a motive of duty nor the regard to the virtue of an action can be the first virtuous motive that renders an action virtuous.

In any case, we face two options, neither outlandish. One is to hold that Hume thought that what matters, in the case of the artificial virtues, is not *why* one acts in conformity to the relevant conventions, but just that one does. The other is to hold that Hume thought that in the case of the artificial virtues the motive still matters, but that the motive of duty (ungrounded though it is) is the motive that matters.⁴⁸

⁴⁸ This interpretation has been advanced by Rachel Cohon in her Hume’s Morality: Feeling and Fabrication, *op. cit.*

In favor of the first option, one might think, is that we count people as having done the just thing (for instance, as having properly returned someone else's property, as having kept their promise, as having given allegiance to their government) without having to investigate their motives. And we regularly approve of their having done so, again without having to investigate their motives. This is in contrast with the natural virtues, it seems, in that we count acts as benevolent, kind, courageous, etc. only in light of the motive with which they were performed. And it might seem we regularly approve of what they have done only in light of their motives. This contrast suggests that with the artificial virtues what matters is the performance, not the motive, and that the performance is what secures the relevant approval. Against this, though, is the observation that our moral *admiration* of the actions in accord with the artificial virtues seems, as in the case of the natural virtues, to depend on the motive. We might be glad that a person did the just thing, regardless of her motives, but our admiration for her, and our regard for her as a virtuous person, depends on the action not being, say, a part of a plan to willfully deceive others, or a reflection merely of fear of punishment. So while it seems we recognize acts as, for instance, the acts demanded by justice, independent of the motive of the agent, we count conformity with those demands as virtuous only in light of the agent's motive.

At the same time, though, if we are in fact able to recognize acts as just, or as instances of keeping one's promise, or as being demanded by allegiance, all without appeal to the motives of the agents, we find some reason in favor of the second option, according to which (contra the circle argument) the motive of duty is the first motive for the artificial (but not the natural) duties. Recall that the circle argument depends on the idea that we cannot identify an action as virtuous, in a particular context, unless and until we find a non-moral motive that would both prompt that action and secure approval from the General Point of View. Having found such a motive, a person might recognize that the motive would lead to a particular action and then perform that action from some other motive (including a motive prompted by recognizing that the action would be virtuous). But absent such a motive, the circle argument supposed, there is no way to identify which actions would be virtuous, and so no way successfully to do a virtuous action because one recognizes it to be virtuous. Yet if, as this interpretation suggests, we can identify which actions are virtuous, in the case of the artificial virtues, without appeal to virtuous motives, the circle argument loses its force, at least with regard to the artificial virtues. And room is left for people to perform such actions from duty – that is, from the recognition that they are virtuous – without having to suppose a non-moral motive to perform such actions.⁴⁹

Of course this stands as an interpretation of Hume only if there is reason to think Hume rejected the circle argument when it came to the artificial virtues. If he did, it was – remarkably – without any comment whatsoever. And he gives no indication either of thinking that, in the case of the artificial virtues, we can identify actions as virtuous without regard to the motives with which it might be performed.⁵⁰

⁴⁹ Indeed, at one point Hume himself claims that “’tis evident we have no motive leading us to the performance of promises, distinct from a sense of duty” (T 3.2.5.6/518-19). But there is reason to think (see above, footnote 33) that what Hume had in mind was that there is no *natural* motive leading us to the performance of promises (other than a sense of duty). And the point Hume seems to be making, in context, is that the circle argument establishes that there must be some non-natural – that is, artificial – motive in play.

⁵⁰ Needless to say, even if Hume actually held tight to the circle argument and so to the undoubted maxim which it supports, one might think that he should have put them aside and that if he had the resulting account of the artificial virtues would be the stronger for it. I think this would be a mistake, given that (as I argue) there is a non-moral yet

This should give pause when it comes to accepting this interpretation. Moreover, as I will argue below, there are in fact non-moral, yet virtuous, motives that play the role Hume argues is needed. And without appealing to them we would be unable correctly to identify and explain the distinctive virtue of justice. This interpretation would force us to count as just only those who act from a sense of duty (or the recognition of the morality of acting in the way they do), which over-moralizes what it takes to be a just person. While one might well act from such a sense, just people, it seems to me, can, and often do, act justly without thought of duty (or morality more generally), in much the same way that people who are benevolent can and do act benevolently without thought of duty (or morality more generally). In the case of benevolence, people count as acting benevolently (and as a benevolent person would) if they act as they do from a concern to help others, which they may do without thought of duty. Similarly, in the case of justice, people count (I will argue) as acting justly (and as a just person would) if they act as they do from a concern to do their share (in a mutually advantageous enterprise), which they may do without thought of duty (or morality more generally). As a consequence, this interpretation would mark, in the wrong place, the difference between virtuous conformity with conventional rules and conformity that is not virtuous, by holding that the former requires acting from a sense of duty. This interpretation is simply too restrictive in what it counts as the motives in light of which we properly admire someone as having (what Hume identifies as) the artificial virtues.

Needless to say, to make good these worries it's important to show that a concern to do one's share (in a mutually advantageous enterprise) is (i) a motive a just person might act on, (ii) a motive in light of which she is properly admired as just, and (iii) a non-moral motive (that is, a motive that need not involve a regard to virtue). In other words, we need to defend it as a non-moral yet virtuous motive that is characteristic of just people.

Before turning to that task, it is worth mentioning two alternative interpretations that take seriously Hume's "undoubted maxim" ("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 3.2.1.7/479)) while retaining the idea that Hume thinks the artificial virtues are genuine virtues. On the first,

The first, non-moral yet virtuous, motive is enlightened self-interest.

This interpretation involves contrasting enlightened self-interest with the self-interest that Hume earlier rules out as not being the motive to justice. The contrast between the two turns on enlightened self-interest being informed by, and available only within, contexts shaped by conventions that make new possibilities available. The enlightenment involved is not just that provided by taking a long and informed view of one's interests, but also by appreciating the benefits collective action makes possible. The idea is that once conventions are in place a narrow and myopic self-interest is reshaped by an appreciation of the individual benefits we each enjoy by being willing participants in mutually advantageous conventions. "There is," Hume notes, "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, 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 3.2.2.13/492).

virtuous motive that is part of a unified and compelling account of what is motivationally characteristic of people who have the artificial virtues.

Lending support to the contrast is the fact that when Hume rejects self-interest as the relevant motive he concentrates on “self-love, *when it acts at its liberty*” (T 3.2.1.10/480, ital. added), whereas the self-interest advanced here is under restraint, albeit by “the very affection itself, by an alteration of its direction.”

In favor of this interpretation are the many places in which Hume emphasizes the extent to which the various conventions he has in mind are in the interest of those who participate in them, and the extent to which a recognition of this in fact leads many people to establish and then conform to the conventions. In fact, there is no question, as a matter of interpretation, that Hume thinks that the relevant conventions owe their existence to peoples’ pursuit of self-interest. Hume claims, for instance, that “the rules of justice are establish’d merely by interest” (T 3.2.2.22/497-8)⁵¹ and he insists that “To the imposition then, and the observance of these rules [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 3.2.2.24/498-500).⁵²

The question is whether the motive of self-interest, enlightened and restrained by an appreciation of the benefits that flow from the conventions, is the first non-moral yet virtuous motive *that secures approval from the general point of view*.⁵³ Might enlightened self-interest be not just the motive that explains why the conventions that are at the heart of the artificial virtues are established but also “the first virtuous motive” which “is requisite to render an action virtuous” (T 3.2.1.4/478)? Here Hume’s arguments against seeing (unenlightened) self-interest as the first, non-moral yet virtuous, motive, appear to stand equally against enlightened self-interest.

The first of these arguments is that self-interest, even enlightened self-interest, will not always lead someone to act as the artificial virtues would require. While it might well prompt one to greater compliance than will be shown by a person driven by self-interest “at its liberty,” there will be cases where a person’s interest is clearly best served by acting contrary to justice. As Hume notes, “’tis easily conceiv’d how a man may impoverish himself by a single instance of integrity, and have reason to wish, that with regard to that single act, the laws of justice were for a moment suspended in the universe” (T 3.2.2.22/497-8). A person motivated by enlightened self-interest will predictably be moved, in such circumstances, to suspend the laws of justice for himself, however they stand in the universe.

At points in the *Treatise*, though, Hume seems not to recognize this, writing as if he thinks conforming to the rules of justice is always, on balance, advantageous. So, for instance, concerning the rules of justice, Hume writes that when “society has become more numerous, and has encreas’d to a tribe or nation” people “do not so readily perceive, that disorder and confusion follow upon every breach of these rules” (T 3.2.2.24/498-500). This suggests that Hume might think that *every*

⁵¹ Hume goes on immediately to stress that the connection between the rules of justice and self-interest is “somewhat singular.” Our interests are not served by each act required by a rule of justice, he notes, but by “the steady prosecution of the rule, and by the peace and order, which it establishes in society (T 3.2.2.22/497-8).”

⁵² “*Thus self-interest is the original motive to the establishment of justice: but a sympathy with public interest is the course of the moral approbation, which attends that virtue*” (T 3.2.2.24/498-500). Hume says something similar concerning the convention of promising. See T 3.2.5.11/522-3.

⁵³ This interpretation is pressed by David Gauthier in “Artificial Virtues and the Sensible Knave” (*op. cit.*), even as he ultimately argues that Hume was committed to an error theory to the effect that our taking justice to impose a duty is a mistake (because, Gauthier argues, the motive required by Hume’s “undoubted maxim” is not actually available).

breach does in fact cause disorder and confusion, and that the problem is just that people have trouble seeing this.⁵⁴

But this suggestion seems wrong, for two reasons. First, it is obvious, and would have been to Hume, that “disorder and confusion” do not “follow upon every breach” of the rules of justice. When discovered, breaches may prompt anger and resentment, but they might well not be discovered, and even when they are, anger and resentment are not the same as disorder and confusion. Neither a single act, nor even several acts, of injustice will work to undermine the whole scheme. Second, Hume clearly recognizes this, observing just one page earlier that “Taking any single act, my justice may be pernicious in every respect [both to my own interest and to the public interest]; and ‘tis only upon the supposition, that others are to imitate my example, that I can be induc’d to embrace that virtue” (T 3.2.2.22/497-8). Yet this is a supposition that Hume knows to be false in certain cases, as when the breach is entirely private.

What is advantageous to all, Hume makes clear, is the “whole plan or scheme” constituted by the rules of justice, which “is requisite, both to the support of society, and the well-being of every individual” (T 3.2.2.22/497-8), not the particular acts required by those rules. Needless to say, the scheme would collapse if people too regularly violated the rules. And peoples’ interests are well advanced by performing the actions that are important to sustaining the “whole plan or scheme.” Yet performing those acts that are crucial to maintaining the relevant conventions is compatible with performing some unjust acts, and a person concerned solely with enlightened self-interest, rather than justice, will presumably take advantage of the opportunities, if she sees them. Hume notes this explicitly in the *Enquiry*: “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.”⁵⁵ Hume does go on to argue that the sensible knave is the true loser, but the loss in question comes only because the sensible knave misses out on the “invaluable enjoyment of a character, with themselves at least, for the acquisition of worthless toys and gewgaws” – a benefit that depends on being someone who values being just and has an “antipathy to treachery and roguery” (EM 9.23/283).⁵⁶

The second argument for thinking self-interest, even enlightened self-interest, is not the relevant motive is that *that* motive is not the motive that secures approval from the general point of view as the motive of justice, even if enlightened self-interest were to lead one always to conform to the

⁵⁴ Gauthier, in fact, makes just this suggestion.

⁵⁵ Gauthier holds that in the *Treatise* Hume mistakenly assumed that the whole scheme of justice depended on each performance, and that this assumption was crucial to his thinking that the first virtuous motive to justice was enlightened self-interest. On this reading, the sensible knave of the *Enquiry* comes on the scene as a revelation and a new challenge to the adequacy of the account advanced in the *Treatise*. Against this proposal stand the passages in the *Treatise* that seem to recognize explicitly that individual acts of injustice may well be in one’s self interest.

⁵⁶ It is worth noting that this observation is compatible with advancing an argument from self-interest for valuing justice and having an antipathy to treachery and roguery, as long as, compared to the alternatives, their contribution to one’s overall interest is greater. But this will be true only if there are already some other interests in place that are better served by cultivating an concern for justice and an antipathy to treachery and roguery. As Hume acknowledges, for a person tempted by the thought “That *honesty is the best policy*... but is liable to many exceptions,” (EM 9.22/282-3) and who holds “that this reasoning much requires an answer, it would be a little difficult to find any which will to him appear satisfactory and convincing” precisely because his actual interests might not be well served by an alteration of his interests (EM 9.23/282-283).

demands of justice. A person who cautiously and successfully pursues her enlightened self-interest may well be admirable as prudent, but the prudence she exhibits does not make her a just person, even if she behaves exactly as a just person would. With the difference in motive comes a difference in the virtue, though virtue may still be there to be found. No matter how enlightened the self-interest, and no matter how successfully it leads one to conform to the rules of justice, the just person is set apart from the prudent person by *why* she acts as she does. And the key difference is that a just person's concern, in acting as she does, is not with her self-interest (enlightened or not) but with the rules or principles established by the mutually advantageous conventions within which she finds herself.

This crucial fact is nicely picked up by those who see Hume as identifying the just person as the person within whom the rules of justice (broadly construed to cover property, promise, allegiance, etc.) serve as an "inviolable law" that resists "any temptation, to be induc'd to violate those principles, which are essential to a man of probity and honour" (T 3.2.2.27/501).⁵⁷ And this brings us to a second and more attractive interpretation, according to which

The first, non-moral yet virtuous, motive is provided by accepting the rules of property, promising, and allegiance, as inviolable laws.

It may well be that self-interest would lead one to become such a person. For it might be that the benefits of "the whole plan or scheme" made possible by certain conventions are (or are thought to be) available only to those embrace the conventions's rules as principles for action. But the (potentially self-interested) motive to become such a person is importantly distinct from the motive one has, if one is such a person. And it is the latter motive that is a plausible candidate for the first, non-moral yet virtuous, motive to justice.⁵⁸

In support of this interpretation are the several places where Hume maintains that a just person will regard the rules of property, promises, allegiance, etc. as "inviolable." To take just one example, Hume writes that "every one, who has any regard to his character, or who intends to live on good terms with mankind, must fix an inviolable law to himself, never, by any temptation, to be induc'd to violate those principles, which are essential to a man of probity and honour" (T 3.2.2.27/501; see also T 3.2.1.15/482-3, T 3.2.6.10/533, T 3.2.10.15/562-3).⁵⁹ The just person is one who treats such

⁵⁷ Stephen Darwall and Don Garrett both emphasize, correctly to my mind, the importance of recognizing that a proper account of justice will find a place for people being motivated by the principles that are "essential to a man of probity and honour."

⁵⁸ Darwall thinks this is the right motive, but argues that Hume's account of the will cannot accommodate the motive, because (according to Darwall) Hume's theory of the will commits Hume to a hedonistic version of psychological egoism that rules out accepting the rules as guides to behavior for their own sake. (See Darwall, *op. cit.*, p. 294.) I think this mistakenly saddles Hume with a moral psychology to which he is – fortunately -- not committed. Consider especially Hume's discussion "Of Self-Love," which appeared as an appendix to An Enquiry Concerning the Principles of Morals for evidence that Hume recognizes people as having concerns beyond their own interests and pleasures. (See Garrett for a nice defense of Hume against Darwall's charge.) Garrett in contrast defends this motive as not just one that Hume can consistently recognize, but as the first virtuous motive Hume thought was characteristic of justice.

⁵⁹ After arguing that self-interest, expressed in the adoption of a policy to comply with the rules of justice, constitutes "the original non-moral motive to justice," Garrett distinguishes that motive from the "first virtuous motive, which bestows a merit" on the actions performed by a just person (p. 271). What is important to keep in mind here is that the "first virtuous motive, which bestows a merit" must itself, by Hume's argument, be a non-moral motive, and a non-moral motive that leads consistently to the performance of just actions, while being the motive we admire as constituting the justice of the person in question. Self-interest, even enlightened and redirected, does not fill that bill. Still, there is no mistaking the fact that Hume recognizes self-interest as the (or, as I think, a) motive that initially leads people to

principles as inviolable. While a person might do so for self-interested reasons, it is worth noting that a concern not to violate such principles without their having acquired it for self-interested reasons. Indeed, as Hume emphasizes, educators and politicians, parents and friends, all seem to work to encourage justice, often (though not always) without recommending it as part of an advantageous plan. Thus parents “are induc’d to inculcate on their children, from their earliest infancy, the principles of probity, and teach them to regard the observance of those rules, by which society is maintain’d, as worthy and honourable, and their violation as base and infamous. By this means the sentiments of honour may take root in their tender minds, and acquire such firmness and solidity, that they may fall little short of those principles, which are the most essential to our natures, and the most deeply radicated in our internal constitution” (T 3.2.2.26/500-1).⁶⁰

So while it may be that some people embrace certain rules as inviolable for self-interested reasons, and so come to have the motive in question, the motive itself does not require consciously embracing such a plan and may well find its firmest place among those acquired the disposition to treat certain rules as inviolable in some other way. Thus, although Hume puts the point saying that a person “*must fix an inviolable law to himself*” (T 3.2.2.27/501, ital. added), the disposition is found in anyone within whom the principles of justice have, in the appropriate way, taken hold, whether or not he has put them there himself.⁶¹ And, in fact, for those who have consciously embraced such a plan out of a regard for their enlightened self-interest might well find themselves wondering whether treating the laws as inviolable is really necessary.⁶²

A person within whom the principles of justice have taken hold (however that happens) are moved by the thought that the something is another’s property, the thought that they have given their word, and the thought that something is required by their country’s law, and not just by self-interest or benevolence. Moreover, in a genuinely just person, these thoughts are not simply among the many motives they may have for acting; they take precedence over “particular views of private or public interest” and stand against “spite and favour” (T 3.2.6.9/531-3). Importantly, such motives can play a preeminent role in guiding a person’s actions, without the person thinking in terms of virtue, or of duty, or of what might garner approval from the General Point of View. They are non-natural motives (in the relevant sense) since they require conventions to give them the content and direction they need to prompt action force, and they are non-moral motives (again, in the relevant sense) since an agent may be moved by a concern with property, promises, the laws of one’s country, without thinking in moral terms at all.⁶³ At the same time, under the right circumstances – when the rules in

establish, and conform to, the conventions at the heart of the artificial virtues. In this sense, self-interest is an original non-moral motive to justice, but not the motive that is required by the circle argument.

⁶⁰ Hume argues that however effective these efforts, to the extent they depend on talk of what is “*honorable or dishonorable, praiseworthy or blameable,*” they presuppose another “cause of the distinction we make betwixt vice and virtue” (T 3.2.2.25/500).

⁶¹ This isn’t the only place where Hume characterizes the restraint as self-imposed: “when any man *imposes on himself* general inflexible rules in his conduct with others, he considers certain objects as their property, which he supposes to be sacred and inviolable” (T 3.2.6.10/533, ital. added).

⁶² This is true even for someone who cares about maintaining his character as a “man of probity and honour” or a concern for “living on good terms with mankind” since these interests might compete with others better served by making an exception to the plan.

⁶³ Hume makes clear that that as he is thinking of things, one might recognize actions as required by justice (or equity) without taking them to be virtuous, though, he thinks we do usually annex the idea of virtue to that of justice. (See footnote 34.) Similarly one might think of actions as required by a convention (or the laws of a country) without thinking of them as virtuous. In the same way, one might have the motives involved in being willing to do one’s share, as specified by a convention, without thinking of doing so as virtuous (or as one’s duty), though of course one might also think that so acting is virtuous (or one’s duty).

play are part of a mutually advantageous convention – being moved by of such thoughts does secure the approval that constitutes them as virtuous.

The caveat, “under the right circumstances,” is crucial, however, and provides reason to worry about over-stressing the idea that a person with the artificial virtues will regard any rules as “sacred and inviolable.” Hume’s talk of inviolable laws, as well as his characterization of people imposing “general inflexible rules” on themselves that lead them to see property as “sacred and inviolable” (T 3.2.6.10/533), risks being misleading when thinking about Hume’s conception of the artificial virtues. Specifically, it risks masking the fact that, on Hume’s view, the rules that “a man of probity and honour” will be concerned to comply will vary according to which conventions (if any) happen to be in place. It is no virtue to stick to the rules in a context where the relevant convention has no point or place (for instance, where others cannot be counted on to do their part). “Suppose,” Hume asks, “... it should be a virtuous man’s fate to fall into the society of ruffians, remote from the protection of laws and government; what conduct must he embrace in that melancholy situations?” He answers: the virtuous man “can have no other expedient than to arm himself, to whomever the sword he seizes, or the buckler, may belong: And his particular regard to justice being no longer of USE to his own safety or that of others, he must consult the dictates of self-preservation alone, without concern for those who no longer merit his care and attention” (EM 3.9/187).⁶⁴ Moreover, a virtuous person will take the “inviolability” of the rules in question to be conditional on the conventions of which they are a part being *mutually* advantageous. In cases where they are not, “a man of probity and honour” need have no disposition at all to comply with the conventions (and may have reason to resist them).

So it is important to see that the first, non-moral yet virtuous, motive for the artificial virtues cannot plausibly be tightly bound up with a disposition to regard some set of rules (either a single set of rules for all situations, or whichever rules are established by the conventions in which people find themselves) as genuinely “sacred and inviolable.” Even the rules of a convention that is mutually advantageous lose their force, on Hume’s account, if the circumstances change in ways that make complying with them no longer an integral part of a mutually advantageous plan or scheme.⁶⁵ As a result, treating the rules of any convention as “sacred and inviolable” will reflect a (presumably vicious) rule fetishism, and not the virtue of justice.

Thus the first, non-moral yet virtuous, motive underlying the artificial virtues is found not in the virtuous person’s taking certain rules as inviolable, but in her commitment to conform to general rules established by conventions, *given* that the conventions are mutually advantageous, and that others are doing their share as well. This commitment requires that one will “never, by any temptation, be induc’d to violate those principles, which are essential to a man of probity and honour.” Yet just what those principles are will vary in significant ways across time and culture. In an important respect, then, on Hume’s view (I am suggesting) the only “inviolable law” recognized by a man of probity and honour is found in the commitment not to violate the general rules of mutually advantageous conventions, *whatever, given their circumstances, such rules turn out to be*. Such a

⁶⁴ Hume emphasizes as well that under conditions of extreme scarcity “the strict laws of justice are suspended...and give place to the stronger motives of necessity and self-preservation.” And he goes on to ask rhetorically “Is it a crime, after a shipwreck, to seize whatever means or instrument of safety one can lay hold of, without regard to former limitations of property?” (EM 3.8/186-7).

⁶⁵ This will happen under two conditions. It might be that others are not willing to do their part. Or, even if they are, it might be that a willingness to comply, as long as others are, stops being *mutually* advantageous – either because mutual compliance does not yield the advantages, or because the advantages do not depend on a mutual willingness to comply.

person is not willing, *unconditionally*, to conform to actual conventional rules that happen to be advantageous. Her willingness is conditional.

Yet when the conditions are met the virtuous person's compliance is unstinting -- "unchangeable by spite and favour, and by particular views of private or public interest" (T 3.2.6.9/531-3).⁶⁶ Indeed, the just person's unwillingness to make exceptions for herself, her commitment to the general rules established by the (mutually advantageous) conventions is central, Hume argues, to securing the benefits the conventions make available. Were people regularly to treat honesty "as the best policy... but one liable to many exceptions," as the sensible knave recommends, our capacity to trust each other, and the benefits such trust makes available, would evaporate. A willingness not to make exceptions in light of "of particular views of private or public interest," assuming others are likewise willing, is essential to securing the benefits justice makes available.

Still, it is a central and attractive feature of Hume's account of the artificial virtues is that people with them are concerned not, in the first instance, with conventional rules (which might change with changing circumstances). They are concerned, rather, with doing their share in mutually advantageous enterprises in which others are likewise willing. Such a person will willingly do her share, provided others are as well, while resisting temptations offered by opportunities to exploit the cooperation of others.⁶⁷ The role of the conventional rules is secondary, though essential: they work to specify what a person's share is, under the circumstances.

This concern – to do one's share (as defined within a mutually advantageous convention), given that others are (or will) as well – gets content, and so direction, only in contexts of specific conventions that are mutually advantageous. So the concern serves as a motive to action only in such contexts: "the conventions of men..." as Hume puts it, "create a new motive" (T 3.2.5.10/522). In the context of mutually advantageous conventions, the concern to do one's share finds expression in actions that conform to the relevant rules in a way that lets us distinguish those who simply conform to the rules (because of fear, or self-interest, or a fetish for the rules) and those who are, in conforming, being just (that is, exercising an artificial virtue). The difference is found in the just person's willingness to do her share (unmoved "by particular views of private or public interest") in mutually advantageous arrangements, provided others are as well. Others might be willing to conform to the rules, but if the conditions of their willingness are not that others too are willing to do their share within conventions that are mutually advantageous, they may be doing what justice requires, but will not be manifesting the virtue of justice.

⁶⁶ This may reasonably be seen as a mitigated inviolability. One could try to reclaim a more robust inviolability by supposing that each of the relevant rules contains in some way "the condition of its authority" (as one might put it). Then the rule would not be, say, 'respect property as defined in the following way,' but, instead, something like 'when property is defined in the following way within a mutually advantageous convention of which one is a member, respect property so defined.' Thought of in this way, the rule itself might be treated as inviolable in a straightforward sense, even if it was, in effect, often irrelevant under actual circumstance. Pursing this, however, ends up bringing one very close to the idea, suggested above, that the one truly inviolable rule in Hume's account is that one not violate the general rules of mutually advantageous conventions, *whatever, given their circumstances, such rules turn out to be.*

⁶⁷ Hume does not wrestle with the important questions of how the justice of an arrangement depends not merely on it being mutually advantageous but on how those advantages are distributed. Nor does he address worries about how the conditions under which a convention may prove mutually advantageous might themselves be exploitative and so not an appropriate baseline for determining what might be just. The latter concern motivates the idea that the relevant advantage must be one compared to how one would fare absent both force and fraud.

People who are concerned to do their share, as specified by mutually advantageous conventions (provided others are willing as well) offers a *motive* (when the conventions are available) that is *artificial* – depending as it does conventions and artifices that reflect the “designs, and projects, and views of men” (T 3.1.2.9/474). At the same time, it is *non-moral*, since a person may be disposed to do her share, in this way, and yet not “annex the idea of virtue” to doing so. Yet it is *virtuous*, since, “a sympathy with public interest” (T 3.2.2.24/498-500) naturally secures approval for this motive from the General Point of View.

The approval, it is worth emphasizing, differs from the approval enjoyed by the natural virtues precisely because the public interest justice serves depends on appreciating “the whole plan or scheme” (T 3.2.2.22/497-8) realized by acting on a reciprocal willingness to constrain the pursuit of interest (public and private) by mutually advantageous rules. It’s not that there is no virtue in the pursuit of interest (public and private) – Hume holds that there is. It’s that one has missed the distinctive virtue of justice until one appreciates that justice works by constraining the pursuit of particular interest by a rules with the following credential: a willingness to comply with them offers mutual advantage.⁶⁸

Conclusion

A striking and important feature of Hume’s account of the artificial virtues – justice, broadly construed -- is that it recognizes, and explains, the huge variety of things just people will do, depending on their circumstances and the conventions that are in place, while capturing the distinctive character of what will be motivating them. Indeed, it offers a unified account of what all such people have in common: A willingness to regulate their behavior by general rules established by shared conventions, *given* that the conventions are mutually advantageous, and that others are doing their share as well. On Hume’s account, and I think plausibly, such a willingness is the non-moral yet virtuous motive that we admire as distinctive of those who are genuinely just.

Hume does not pursue this idea, but it is worth noting that his account of the artificial virtues fits well not just with our admiring as just those who willingly regulate their behavior by the appropriate rules, but with our thinking that, in doing so, they should be sensitive to whether the conventions that establish the rules are advantageous to all who are subject to them (this is a crucial part of what makes the rules appropriate). In fact, there is room to argue, though again Hume does not pursue this idea, that a just person will be especially sensitive to whether the rules in place are advantageous to the most vulnerable, since the conventions are most likely to lose their authority (as established by

⁶⁸ Any account of justice that treats its demands as dependent on convention faces two familiar worries. The first is that it will be excessively conservative – recommending that a just person embrace whatever conventions happen to be in place. The second is that it will be unsatisfyingly relativistic – leaving out of the picture any sense of there being standards for justice independent of the particular practices one finds in different societies. I hope that it is clear that Hume’s account resists quick versions of these worries. First, Hume’s defense of conventions has at its core an insistence that the relevant conventions are only those that are mutually advantageous. So while a convention must be in place for justice to require compliance, the mere fact that a convention is in place does not mean justice requires compliance with it. (It is true, thought, that Hume’s theory does imply that if a mutually advantageous convention is in place, justice requires compliance with it, assuming one has benefitted by the compliance of others, even if there are other conventions in prospect that one would prefer or that offer, say, a more equal distribution of advantages.) Second, the non-moral yet virtuous motive to justice – the conditional willingness to comply with mutually advantageous conventions, assuming others will as well – in effect embodies a standard of justice that is independent of particular practices found in different societies, by requiring that these practices, in order to establish duties of justice, must provide benefits to all who are subject to them.

what will secure approval from the General Point of View) by failing to address the interests of the vulnerable.⁶⁹

⁶⁹ There is, as I say, room to make this argument. The details of how such an argument would go, though, are complicated by the fact that to fall within the ambit of justice at all, people must be such that their compliance with conventions would be advantageous to others. Those whose don't meet this requirement will still be proper objects of moral concern, as long as they are sentient, Hume emphasizes, but, on his account, that concern will not be a matter of justice. So the vulnerable who are covered by this sort of argument may well not be the most vulnerable people. As a result, part of the complication involved in arguing that justice requires being sensitive especially to the in which a convention affects the vulnerable is distinguishing that demand of justice, with its focus on those who might participate in the conventions in a way that benefits others, from the demands of humanity that require a concern for those who See footnote 40.