

Justice and the Tendency towards Good: The Role of Custom in Hume's Theory of

Moral Motivation

James Chamberlain
University of Nottingham

Forthcoming in *Hume Studies*
Please cite published version

Abstract: Given the importance of sympathetic pleasures within Hume's account of approval and moral motivation, why does Hume think we feel obliged to act justly on those occasions when we know that doing so will benefit nobody? I argue that Hume uses the case of justice as evidence for a key claim regarding all virtues. Hume does not think we approve of token virtuous actions, whether natural or artificial, because they cause or aim to cause happiness in others. It is sufficient for the action to be of a type which has "a tendency to the public good" for us to feel approval of it, and to be motivated to perform it. Once we are aware that just actions *tend* to cause happiness, we approve of all just actions, even token actions which cause more unhappiness than happiness.

1. Introduction

In the Book of his *Treatise* devoted to morality, Hume gives more consideration to what he calls "artificial virtues" than to natural virtues (T 3.3.1.9; SBN 577).¹ Artificial virtues rely on conventions which have arisen for non-moral reasons, but which come to be understood as beneficial to society so that, as a consequence, we come to see adherence to these conventions as being morally obligatory. The main such virtue which Hume discusses is

James Chamberlain, Department of Philosophy, The University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom. E-mail: james.chamberlain1@nottingham.ac.uk

“justice”, by which he means primarily a respect for the conventions concerning property rights and property transfer, such as the convention to pay back loans of money (T 3.2.1.1; SBN 477). Despite Hume’s attention on justice, there is much that is puzzling about his understanding of this virtue, particularly about his understanding of our motivation to act justly. Why do we feel obliged to act justly or honestly on those occasions when doing so will benefit nobody? Why should we, to use Hume’s example, feel duty-bound to return a loan of money to a “miser” or “profligate debauchee” (T 3.2.1.13; SBN 482) when to do so will cause only harm to all concerned?

Hume clearly recognizes that it is possible to feel such motivations. In the second *Enquiry*, he suggests that only a “sensible knave” would decide to break the rules of justice, even in such situations (EPM 9.22; SBN 282-3). Hume believes both that most of us are motivated to act justly in such cases and that we disapprove of those who do not. Yet why should we consider it vicious to break the rules on occasions where no one is harmed? Why are we not sensible knaves?

In a key passage, Hume focuses on two potential motives for justice. These are “self-interest”, which he calls the “original motive to the establishment of justice”, and “a sympathy with public interest”, which he describes as “the source of the moral approbation, which attends that virtue” (T 3.2.2.24; SBN 499-500, italics omitted). These appear to be the *only* potential motives to perform just actions, according to Hume. Yet he allows that some just actions are both to the detriment of the actor’s interests *and* such that “the public is a real sufferer” (T 3.2.2.22; SBN 497). What, then, is Hume’s explanation of the motivation to act justly in cases like this?

I will refer to this explanatory requirement as the problem of “formally just” actions. A formally just action is one which is morally obligatory (insofar as it conforms to the rules of justice), but which the agent considering the action believes will cause no happiness to

anyone. Indeed, formally just actions may be believed by the agent to cause only harm to all concerned.² The problem is of explaining why, in the context of Hume's wider theory of moral motivations, agents might be motivated to perform formally just actions, and to disapprove of those who do not.

I proceed as follows. In Section 2, I discuss the context of Hume's treatment of the problem of formally just actions, in which it is closely related to his discussion of the original motive for justice. Section 3 surveys existing interpretations of Hume's response to this problem and suggests that they each face insurmountable objections. In Section 4, I consider Hume's distinction between natural and artificial virtues and argue that, for both kinds of virtue, he claims that people initially perform the relevant actions purely from non-moral motives, and that they then approve of these motives and actions when they see the benefits which these typically produce. The relevant non-moral, or "original", motive for justice is self-interest. In Section 5, I argue that Hume's response to the problem of formally just actions is that we are now so accustomed to the pleasing effects of justice that we habitually approve of any self-interested or dutiful motive to act justly, and of any action which typically results from such a motive, regardless of our beliefs about their particular effects. Therefore, anyone faced with the prospect of performing a formally just action will approve of the idea of doing so. Approval is a form of pleasure, for Hume, and pleasures produce desires, so this explains why we are motivated to perform formally just actions.

2. Moral Obligation and Justice

When Hume first discusses the virtue of justice, he focuses on what appears to be a fundamental problem with the notion of duty towards justice. Typically, when we act justly, we are motivated by a sense of duty. Indeed, Hume seems to claim that *whenever* we act justly, we do so purely from a sense of moral duty. We "have naturally no real or universal

motive for observing the laws of equity, but the very equity and merit of that observance” (T 3.2.1.17; SBN 483). However, Hume acknowledges that this claim appears to conflict with another proposition which he holds to be true, that “no action can be equitable or meritorious, where it cannot arise from some separate motive” (T 3.2.1.17; SBN 483). By a “separate motive”, Hume means a motive distinct from the sense of duty. His point is that no action can be morally obligatory unless we have some motive other than a sense of duty to perform it.

This latter claim is a consequence of Hume’s theory that we take an action to be virtuous because we feel approval towards the motive or desire which led to that action. If I decide to be just, benevolent or in any way virtuous, my motivation cannot *only* be the dutiful desire to be ‘do the right thing’, or to ‘be virtuous’. This would be “to reason in a circle” (T 3.2.1.4; SBN 478), because, for Hume, ‘virtuous’ and ‘approvable’ are the same thing. If I decide to be virtuous, then I decide to do that which is approved, but what is approved is simply what is virtuous. Hume derives from this 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; SBN 479).

This antecedent, non-dutiful motive to perform any virtuous action is what Hume calls the “original motive” for that virtue (T 3.2.1.13; SBN 482). It is easy to see what the original motives are for natural virtues. Hume gives the example of good parenting. We are generally not motivated to act in our child’s interests because we think we ought to, but because we instinctively want to do so. Hume calls this instinct “natural affection” and claims that it is very clearly “a motive to the action distinct from a sense of duty” (T 3.2.1.5; SBN 478). Hume allows that someone may decide to help their child from a sense of duty, but only because natural affection is already “common in human nature” (T 3.2.1.8; SBN 479). If I normally delight in my child’s happiness, then on an occasion when I feel disinclined to act to please my child I will be pained by this disinclination. This may lead me to act “from a

certain sense of duty” (T 3.2.1.8; SBN 479). But this feeling of duty is, as Hume says with regard to benevolence, a “secondary consideration” (T 3.2.1.6; SBN 478). If I did not realize that good parenting tends to cause happiness, then I would feel no duty to be a good parent.

The problem regarding justice is that it seems to have no original motive. I will call this the “original motive” problem. Hume allows that there is a strong intuition that we are only ever motivated to be just from a sense of duty, and that there is no other obvious motive for just actions in many cases. Yet he also seems to claim that we could only feel duty-bound to be just if we also possess a distinct original motive towards justice. As Hume acknowledges, this appears to be “sophistry” (T 3.2.1.17; SBN 483).

If we can understand how Hume resolves the original motive problem, such that he explains how we come to feel duty-bound to act justly in all cases, then this will resolve the problem of formally just actions. However, this is no easy matter. There is a formidable range of commentators, including Árdal (*Passion and Value in Hume’s Treatise*), Ayer (*Hume*), Harrison (*Hume’s Theory of Justice*) and Gauthier (“Artificial Virtues and the Sensible Knave”), who think that Hume’s account of our motivation for our own, and approval of other people’s, just actions is ultimately flawed. Baron (“Hume’s Noble Lie”, 273) goes so far as to suggest that Hume tells a “noble lie”, and that he ultimately recognizes that we have no moral reasons to perform many acts of justice, but refuses to state this, for fear of discouraging just actions in his readers.

There are however three potential options which Hume has at his disposal, each of which will allow him to consistently categorize justice as a virtue and explain our regard for formally just actions. First, he could point to his distinction between useful and agreeable virtues (T 3.3.1.30; SBN 590). Although most virtues please by sympathy, Hume allows that some virtues please immediately, rather than because we are pleased that they cause happiness in others. Perhaps justice is like wit, which pleases us directly in this way. If acting

justly is always immediately agreeable both in its performance and on witnessing it, then we need look no further for reasons why we are motivated to be just or why we approve of just actions.³ Second, Hume could argue that, despite appearances in some cases, we believe that all just actions *do* somehow directly produce sympathetic pleasures or prevent sympathetic pains. For example, on Stroud's interpretation Hume thinks we believe that any omission of a just action would cause the institution of justice to "collapse" (*Hume*, 214). If this were the case, then one's motive to perform any just action would result, at least in part, from one's desire not to harm society at large. The third option for Hume is to argue that all just actions produce pleasures, or avoid pains, which are ultimately but *indirectly* derived from our sympathies with others. We will see that this is the most fruitful available option for Hume, and there are several interpretations of his theory which take this line.⁴ In the next section, I will consider all three options, and argue that none of them are defensible as they currently stand.

3. Potential Solutions to the Problem of Formally Just Actions

In what follows, I will consider each of the three options in turn. I will argue that the final option is the most plausible, but that resolving the problem of formally just actions in this way points to a reinterpretation of Hume's theory of approval generally.

3.1. Is justice an immediately agreeable virtue?

Hume tells us that some character traits are considered virtuous because the actions in which they result are immediately agreeable, rather than because they cause happiness in those around us (T 3.3.4.8; SBN 611). While formally just actions are not useful to anyone, they may perhaps be intrinsically pleasing to perform and to witness. For example, Krause suggests that we are taught that acting justly is inherently pleasing, by politicians and parents

who “make us regard the observance of the rules by which society is maintained ‘as worthy and honourable’ in their own right, and their violation as ‘base and infamous” (“Hume and the (False) Luster of Justice”, 644). However, Krause also believes that Hume fails to articulate a persuasive account of the motive to justice, because the potential benefits of acting unjustly in such situations are so great and because he has shown no way of assessing these potential pleasures against the pleasure of acting justly (“Hume and the (False) Luster of Justice”, 645). An additional problem is that Hume frequently stresses the usefulness of justice, and asserts plainly that “*a sympathy with public interest is the source of the moral approbation, which attends that virtue*” (T 3.2.2.24; SBN 499).

3.2. Do all just actions cause sympathetic pleasures or prevent sympathetic pains?

Perhaps Hume believes that *all* just actions produce happiness to those around us, despite occasional appearances to the contrary – in other words, that there are no formally just actions. For this to be the case, Hume must believe that justice is a fragile enough institution to be threatened if any token just action is disregarded. If this were true, then we would expect painful consequences, or at least a non-negligible probability of painful consequences, from any omission of a just action. There are admittedly some indications that Hume holds such a view. He offers a quasi-historical account of how we came to have rules of justice, which suggests that the levels of cooperation we require as human beings in order to live comfortably on limited resources require us to develop such rules, initially from self-interest as a pragmatic measure. As Cohon summarizes it, we realize that we can “invent rules attaching goods to individuals, and experiences with the fragility of small societies teaches us that it is in our interest to conform to these rules” (*Hume’s Morality*, 173). During this quasi-historical account, Hume suggests that, although some just actions may cause unhappiness to all concerned, “every individual person must find himself a gainer, on ballancing the account;

since, without justice, society must immediately dissolve, and every one must fall into that savage and solitary condition, which is infinitely worse than the worst situation that can possibly be suppos'd in society" (T 3.2.2.22; SBN 497).

However, it does not seem plausible that anyone faced with the choice of repaying a miser will believe that society will collapse if they fail to do so. If that is Hume's argument, then he is providing only a forced choice, as Stroud says, "between a world in which everyone including himself is always just, and a world in which he and everyone else is always unjust", and so one where a just society collapses (*Hume*, 210). I do not think we should consider this to be Hume's argument in the quoted passage. Here, Hume is not considering the reasons why we in a *modern* society would feel motivated to act justly. He is arguing that, in primitive and small societies, the rules of justice are developed and obeyed from simple self-interest. It is only *later* that justice becomes approved of, and is therefore regarded as a virtue.

In the paragraph in question, Hume is still addressing the origins of the convention in a small society. His point is simply that, as soon as people in a small society design property rules, they must realize that the rules cannot allow of exception if they are to survive. In a small community not yet used to acting according to such rules, this seems plausible. It does not follow that in larger societies the same motive will hold. As we have seen, Hume acknowledges that we may firmly believe that breaching the rules of justice will have no negative effects but still disapprove of someone who does breach the rules, and be disinclined to breach them ourselves. Further, Hume has not yet discussed our *moral* motive to perform just actions at all, for he is discussing a point in time where that has not yet arisen. It is in the following paragraph that Hume asks "[w]hy we annex the idea of virtue to justice, and of vice to injustice" (T 3.2.2.23; SBN 498, Hume's italics).

If Hume does not think that we (in our current large-society context) would expect

society to collapse should we omit particular acts of justice, then there is no obvious reason to think that all just actions must cause more happiness than their omission. This brings us to the third possible solution for Hume, which is to claim that at least some acts of justice are motivated by the expectation of other kinds of pleasure than those derived directly from sympathy with others.

3.3. Do formally just actions cause pleasures not immediately derived via sympathy?

If justice – and other artificial virtues – can be shown to have a motive which is not directly related to sympathy, then this might potentially explain our motivation to perform formally just actions. The most commonly suggested motive here is self-interest, or a minimally extended version thereof, because Hume explicitly states that the rules of justice are initially developed and pursued out of mutual self-interest and a concern for the well-being of those dear to us: what Hume calls “*confined generosity*” (T 3.2.2.18; SBN 495). In this vein, Baier (*A Progress of Sentiments*), Bricke (*Mind and Morality*) and Stroud (*Hume*) interpret Hume as holding that we have self-interested reasons for universal compliance to the rules of justice. It is for this reason that Stroud, who thinks Hume’s position is untenable, claims that Hume makes the implausible argument that society could collapse after any one token act of justice is omitted. Bricke similarly takes Hume to believe that it is *always* in one’s narrow interests to comply with the rules of justice (*Mind and Morality*, 218). This thesis is so implausible that Bricke is surprised at what he takes to be Hume’s “untrammelled confidence” in it (*Mind and Morality*, 216).

The main concern with this general interpretation, however, is again that Hume only discusses purely self-interested motivations towards just actions as occurring in the *early stages* of the development of the motive to act justly. It is not until after he has concluded his account of the origins of justice that he considers why we are *morally* motivated to act justly.

He is careful here to say that the reader must wait for the full answer to this, which is in “the *third* part of this book” (T 3.2.2.23; SBN 498). There is therefore no reason to think that Hume attempts to explain our motivation to act justly purely on self-interest. The motive of self-interest is insufficiently moralized for it to be the motive Hume describes.

It may however be that just actions are pleasing because they are *generally* strongly associated with sympathetic pleasures, and that this is sufficient for us to approve even of rare token actions which will not produce such pleasures. This is, I think, the only plausible solution, but the question is how this association causes approval in such situations.

Interpretations along these lines often take Hume to have a complex story about our approval of justice. For example, Cohon argues that there are three stages to the formation of an artificial virtue such as justice. The first stage is straightforward, as we create and follow the rules from enlightened self-interest (*Hume’s Morality*, 173). Then there are two further stages as societies grow, which lead to this conformity becoming a moral practice. First, the mechanism of sympathy naturally makes everyone approve of other people’s rule-following, because of the happiness which such rule-following causes and the misery which it prevents. Then there is a second artifice, this time performed by politicians and parents, which causes us to approve of honest characters as well as of honest actions. It is only after this second artifice that we are morally motivated to perform formally just actions, by the resulting “enhanced moral sentiment” (*Hume’s Morality*, 175).

By Cohon’s interpretation, Hume allows that we initially approve *purely* of just actions, regardless of motive, so long as they conform to the rules (*Hume’s Morality*, 174, footnote 10). Harris thinks Hume allows this for all cases of our approval of justice (“Hume on the Moral Obligation to Justice”). However, Hume argues that, for any virtue, we *initially* approve only of motives: “Actions are at first only consider’d as signs of motives” (T 3.2.1.8; SBN 479). Over time, the close association between virtuous motives and the actions which

they typically produce is likely to cause us to approve of the actions themselves. Eventually, we may sometimes lack a token virtuous motive but nevertheless perform the associated action, “merely out of regard to its moral obligation” (T 3.2.1.8; SBN 479). However, although Hume therefore allows that we *can* act justly because we feel approbation towards the idea of a just action, this can only occur once we habitually associate such actions with the relevant motive for justice: the fundamental source of this approbation.

Perhaps, however, we find justice generally pleasing because we approve of the *motive* to unquestioningly follow the rules, rather than of the rule-following actions themselves. This would be the case if the just person’s motive is to regulate her conduct “by rules she regards as authoritative”, as Darwall suggests (“Motive and Obligation in Hume’s Ethics”, 440). While Darwall takes this to be Hume’s account, he also argues that it is incompatible with Hume’s theory of the will. Darwall understands this theory to stipulate that all just actions must be motivated by desires or aversions, caused by beliefs that pleasures will result or pains be avoided by so acting (“Motive and Obligation in Hume’s Ethics”, 423). However, Darwall argues that the desire to follow a self-imposed rule would stem from no such belief, so that Hume’s theory of justice can only be accepted if his “official theory of will is jettisoned” (“Motive and Obligation in Hume’s Ethics”, 420).

Besser-Jones and Garrett both endorse Darwall’s rule-regulation interpretation, but attempt to exonerate Hume from inconsistency. Besser-Jones argues that the motive to regulate one’s conduct by the rules of justice can be explained by reference to agents’ desires to “develop a good reputation and so become proud of their character” (“The Role of Justice in Hume’s Theory of Psychological Development”, 272). Because we desire to maintain our reputation and character, both in the eyes of others and to ourselves, and because we are proud of our consistent adherence to justice, we are motivated to perform *all* just actions by the belief that this is necessary to preserve our pride and reputation. Therefore, Besser-Jones

argues that a rule-regulation account is compatible with Darwall's interpretation of Hume's "theory of the will" (Besser-Jones, "The Role of Justice in Hume's Theory of Psychological Development", 262).

Garrett similarly aims to defend the rule-regulation account against Darwall's charge of inconsistency ("The First Motive to Justice"; *Hume*). He argues that Hume allows for motivating passions to be caused by beliefs about the benefits and risks of general *policies*. When self-interest motivates people to create the convention of justice, it also "gives rise to a new motive that could not have existed before: the desire and standing disposition to govern or *regulate* one's behavior by the rules of property" (Garrett. *Hume*, 267). As we see the benefits of the entire scheme of justice, and as we similarly see that it is one which rests on everyone's observance of it, we come to approve of those who follow the rules unthinkingly, because we believe that such unthinking adherence to the rules is the best way of ensuring the efficiency of the scheme overall. "In coming to regulate one's behavior in this way, one undertakes to refrain from weighing up the specific advantages and disadvantages of following the rules of property before acting in each individual case, for such weighing would often lead to violations" (*Hume*, 267).

One worry here is that, while Garrett can explain why we might rarely *recognize* the harmful effects of certain acts of justice before we perform them, it is unclear that he can explain why we would feel duty-bound to perform a rule-following action in cases where we *do*, perhaps inadvertently, come to believe that only harmful consequences will result. However, there is a more serious problem for both Garrett's and Besser-Jones's interpretations, and indeed for all rule-regulation accounts. By any such account, Hume is understood to suggest a moral motive for justice which is quite unlike that for the natural virtues. It is essential to such interpretations that, once the rules of justice are established, just persons form a resolution to follow them unswervingly. This resolution is both the motive to

perform just actions and the mental state we approve of in the just person. Of course, no such motive is required to explain our approval of the natural virtues – benevolence does not necessarily involve a *resolution* to consistently help others. Rather, any token desire to help others simply causes approbation. However, Hume is thought to require that we *resolve* to be consistently just, because this is deemed necessary for him to explain why we approve of *all* just actions, including formally just actions.

However, Hume argues that we can explain our approval of just actions in the same way that we explain our approval of *naturally* virtuous actions. Once the rules of justice have been established, “the sense of morality in the observance of these rules follows *naturally*, and of itself” (T 3.2.6.11; SBN 533, Hume’s italics). Admittedly, as Besser-Jones stresses, Hume does also suggest a “new artifice” by which “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, SBN 533). However, he minimizes the importance of this. He later emphasizes, again, that once we have developed the relevant conventions and rules, justice is “*naturally* attended with a strong sentiment of morals; which can proceed from nothing but our sympathy with the interests of society” (T 3.3.1.12, SBN 579, Hume’s italics).

If Hume believes, not only that our *approval* of just actions develops “naturally”, but also that we are typically *motivated* by this approval, then he clearly means to explain our motive for justice in the same way that he explains our moral motivation to perform naturally virtuous actions. As desires to regulate oneself by rules are *not* required to explain the natural virtues, Hume cannot require any such motives for his explanation of justice.

We cannot simply assume that Hume thinks we are motivated to be just by our approval of justice.⁵ Árdal (*Passion and Value in Hume's Treatise*) and Baier (*A Progress of Sentiments*) understand Hume to argue that, although we *approve* of justice because it serves

the public good, we remain *motivated* to act justly purely from self-interest. Certainly, Hume denies that we act from any “love of mankind, merely as such” (T 3.2.1.12; SBN 481). This is because no such passion exists: we do not *instinctively* love strangers. He also denies that we look “so far as the public interest” when we act justly (T 3.2.1.11; SBN 481). We care little for the abstract idea of the public interest, and we do not perform just actions because we reason that they will benefit society. Any motive so formed would be too weak to overcome our self-interested desires, as justice frequently demands.

However, Hume does *not* conclude that our sole motive for justice is self-interest. He argues that our concern for the public good comes, not from instinct or reason, but from sympathy with the happiness of others (T 3.2.1.12; SBN 481-2). Once we are accustomed to justice, we are pleased – “by sympathy” – that it benefits the public, and we therefore come to approve of justice (T 3.3.1.9; SBN 577). Of course, this cannot explain why we *invented* justice, and so Hume argues that our original motive for justice was self-interest, whereas our approval developed later (T 3.2.2.24, SBN 499).

Hume believes that we, in our “civiliz’d state”, *are* motivated by a “sense of duty and obligation” towards justice (T 3.2.1.9; SBN 479). We follow its rules *because* of our sense of duty, which gives us a “regard to justice, and abhorrence of villainy and knavery” (T 3.2.1.9; SBN 479). Wherever we act from duty, we do so because our human nature possesses “some distinct principles, which are capable of producing the action, and whose moral beauty renders the action meritorious” (T 3.2.1.8; SBN 479). By this, Hume means that to be motivated by duty is to be motivated by our approval of the idea of an action. Ideas of acting justly motivate via approbation, while ideas of acting unjustly repel us via disapprobation: “a considerable motive to virtue”, for those of us who are not knaves (EPM 9.23, SBN 283). Indeed, for non-knaves, these sentiments constitute a sense of moral obligation sufficient to motivate even formally just actions. We feel such sentiments because the *system* of justice

tends to cause happiness, with which we sympathize (T 3.3.1.9, SBN 577). The question, therefore, is how the tendency of justice to cause happiness leads, *naturally* and via sympathy, to our approbation of, and so motivation for, *all* just acts.

In the next section I will argue that Hume believes that we feel morally obliged to perform any action wherever we *generally* associate such actions with a “tendency to the public good” (T 3.3.1.12; SBN 580). This interpretation certainly resembles the rule-regulation views of Darwall, Besser-Jones and, especially, Garrett, but there are two key differences. First, I do not take Hume to require that we must *resolve* to be consistently just in order to be consistently just. Second, I do not take Hume to argue that one’s desire to perform a formally just action stems from any *belief* about beneficial consequences, whether of the token action or of justice more generally. Our awareness of justice’s tendency to cause public good is sufficient to produce moral desires to act justly, even where we believe that a token just action will produce only harm. To understand Hume’s argument to this effect, we must first look to his account of justice and its place in his wider account of morality.

4. Instinct, Artifice and Hume’s Arguments about Justice

Consider the three-part structure of the third Book in Hume’s *Treatise*, ‘Of Morals’. The first part is concerned to argue for the essential role of passions, over and above reason, in morality. The second – and largest – part is devoted to examining the artificial virtues. Here Hume argues, among other things, that we are motivated to perform and approve of formally just actions. The third part is ostensibly an account of the natural virtues, but in fact draws heavily on the prior discussion of the artificial virtues and discusses aspects of morality relevant to both kinds. Hume suggests several times in this section that his discussion of our approval of artificial virtues has important implications for his theory of our approval of natural virtues (T 3.3.1.10, 3.3.1.12, 3.3.6.1; SBN 577-8, 579, 618).

I believe that Hume's aim in discussing the artificial virtues *before* the natural virtues is to demonstrate that a general association between an action kind and the happiness of others is sufficient to cause approval of *any* token action of that kind. Because Hume thinks this is easier to show for the artificial virtues than for the natural virtues, he discusses the artificial virtues first, and then applies this thesis to the natural virtues: "We have happily attain'd experiments in the artificial virtues, where the tendency of qualities to the good of society, is the *sole* cause of our approbation, without any suspicion of the concurrence of another principle" (T 3.3.1.10; SBN 578, Hume's italics). Hume could readily have said that the *good* caused to society by the artificial virtues, or our *belief* in this good, is the sole cause of our approbation towards them, had that been his meaning. Instead, he explicitly refers to the *tendency* of these qualities as being the cause.

In the final part of his Book on morals, Hume describes the cause of approval as being the tendency to cause happiness (or that of disapproval as the tendency to cause unhappiness) on no fewer than nineteen occasions.⁶ At no point does Hume suggest that we can only be morally motivated to perform an action where we hold certain *beliefs* about the consequences of doing so. Of course, this is not conclusive. No doubt one could read a sentence like "qualities acquire our approbation, because of their tendency to the good of mankind" (T 3.3.1.10; SBN 578) as meaning that we approve of qualities on token occasions when we *believe* they will cause some good. However, read literally, Hume is saying that we approve of quality types simply because they *tend* to cause good. With this in mind, the frequency and importance of the word 'tendency' in this final part of the Book cannot be denied.

There is a good reason for Hume to begin arguing his case by discussing the artificial virtues, rather than the natural virtues. It is only with virtues like justice that we are likely to observe that some token actions are considered morally obligatory even where we do not expect them to cause happiness to anyone. One would rarely, if ever, be faced with a situation

where one feels duty-bound to act kindly but where one knows that only unhappiness will result. It is simply not in the nature of kindness, unlike that of justice, for this to be a realistic possibility. Therefore, Hume begins his discussion of virtues by analyzing the artificial virtues, because these provide crucial evidence for his thesis. If he can show that we approve of artificial virtues because they are *generally* associated with causing sympathetic pleasures, then he can claim the same form of general association as the cause of our approval of the natural virtues.

This is why Hume's discussion of justice focusses on the original motive problem, for it can only be this *original* motive which we come to associate with causing sympathetic pleasures, and which thereby causes us to approve of *all* just actions. The distinction between the natural and artificial virtues is, as we have seen, predicated on the former having *instinctive* 'original motives' which the latter lack. These original motives cause us to act in ways which, as we swiftly realize, tend to cause happiness in others. Taylor perspicaciously notes that "in his sketch of our pre-just moral psychology, the motives that Hume focuses on – sexual appetite, affection for children, limited benevolence, interest and resentment – are ones he characterizes in Book II as 'calm desires and tendencies'" ("Justice and the Foundations of Social Morality in Hume's Treatise", 10). These kinds of motive are unusual in that they cause us to act without the expectation of pleasure, but simply "arise from a natural impulse or instinct, which is perfectly unaccountable" (T 2.3.9.8; SBN 439). Some of these instinctive motives tend to cause happiness to others, and these motives constitute the natural virtues.

Hume explicitly describes the instinctive aspect of natural virtues in "Of The Original Contract", where he says that people act in such ways because they are "impelled by a natural instinct or immediate propensity, which operates on them, independent of all ideas of obligation, and of all views, either to public or private utility" (Original Contract, 479).⁷ This

is contrasted with artificial virtues, which are “such as are not supported by any original instinct of nature, but are performed entirely from a sense of obligation, when we consider the necessities of human society, and the impossibility of supporting it, if these duties were neglected” (Original Contract, 480).

Garrett argues that Hume does not mean *moral* duty by “obligation” here (“The First Motive to Justice”, 263). He points to Hume’s discussion of justice in the *Treatise*, where Hume suggests that we are initially motivated to follow the rules of justice by self-interest, and where he calls this motivation a “*natural* obligation” (T 3.2.2.23; SBN 498). However, Hume does not mention the concept of natural obligation in *Of the Original Contract*. Equally, he would not have assumed that his reader would have read the *Treatise*, given that he had disowned it by the time of writing this essay. Furthermore, it is clear from Hume’s description of natural virtues that he *is* referring to a specifically moral obligation. The instincts which cause us to perform natural virtues are contrasted with motives arising from ideas of obligation: an obligation which arises later when we come to pay natural virtues “the just tribute of moral approbation and esteem” (Original Contract, 479).

By the end of part two of Book 3 of the *Treatise*, then, Hume has divided the virtues into those which consist of (motives for) action types that we are instinctively inclined to perform and those which consist of (motives for) action types that we are not instinctively inclined to perform. These latter virtues come into being because of human ingenuity and contrivance, and so are called ‘artificial’. The natural and artificial virtues have this in common: in neither case do we initially perform the relevant actions with any idea of morality in mind. We are initially motivated to perform natural virtues by instinct, and we are initially motivated to conform to the rules of justice from self-interest. However, in both cases, because acting in these ways tend to produce happiness in others, all who witness such actions are frequently pleased by sympathy, and become accustomed to feeling pleased in

this way. This leads us to approve of the motives to perform these actions – whether instinctive or self-interested – and to categorize them as virtues. Eventually, we approve of the actions which typically result from these motives as well, as Hume explains at *Treatise* 3.2.1.8 (SBN 479). From this point on, we will feel obliged to perform such actions even when the motives are absent. We will therefore perform just actions even when we have *no* self-interested reason to do so.

Hume is now able to make his central claim about the causes of approval. If we naturally come to approve of justice because the motives and actions involved *tend* to cause sympathetic pleasures in us, then such a tendency is sufficient to cause approval of any trait. Therefore, Hume’s answer to the problem of formally just actions is the same as his answer to “Virtue in rags”: the problem of explaining why we approve of those who aim to be virtuous but are prevented by circumstance (T 3.3.1.19; SBN 584). In both cases, Hume’s answer is that “where any object, in all its parts, is *fitted* to attain any agreeable end, it *naturally* gives us pleasure, and is esteem’d beautiful, even tho’ some external circumstances be wanting to render it altogether effectual” (T 3.3.1.20; SBN 584, my italics). This is not a concession to a difficult objection, nor does it require any amendment to Hume’s theory of approval. Indeed, it *is* Hume’s theory of approval. In order to fully understand *why* Hume takes this line, we must look to his accounts of causal reasoning and of the moral sentiments.

5. Custom and Sentiment: The Causes of Moral Passions

In Section 3, we saw that Hume thinks non-knaves are motivated to perform formally just actions by a sense of duty, where this is understood to mean that they are motivated purely because they approve of the idea of performing the action and disapprove of the idea of not performing the action. In Section 4, we saw that Hume thinks we approve of any token motive or action of a type which generally pleases us by sympathy. Therefore, we should take it that

we are motivated to perform formally just actions purely by our moral sentiments, which we feel towards formally just actions due to the general tendency of just actions to please us via sympathy. The remaining puzzle is why, on Hume's view of moral psychology, we are motivated by moral sentiments to perform *any* actions which we believe will cause only unhappiness for us and others.

For Hume, any motive is an "impulse of passion" (T 2.3.3.4; SBN 415). Some motivating passions are caused directly by instinct (T 2.3.9.8; SBN 439). The remainder are caused by the "perception of pain and pleasure", which produces aversions for painful objects and desires for pleasurable ones (T 1.3.10.2; SBN 118). This form of motivation may occur when one experiences a *feeling* of pain or pleasure, as when one touches a painfully hot surface. Wherever we *feel* a sympathetic pleasure from the performance of some action, we will be directly motivated to continue this action by our pleasure at the happiness of others. Equally, approbation is a form of pleasure, and disapprobation a form of pain, so they too are such as to produce desires and aversions (T 3.1.2.3; SBN 471). Indeed, it is precisely their "influence on human passions and actions" which demonstrates to Hume that they are passions, rather than the conclusions of reason (T 3.1.1.5; SBN 457). Therefore, where we approve of the action we are performing, we will also be *morally* motivated to continue it.

Our question, however, is why Hume thinks we are morally motivated to perform some *future* actions, where we believe these will cause only unhappiness. Here, the relevant form of motivation is that caused by the "prospect of pain or pleasure", where the painful or pleasurable object appears only as an idea (T 2.3.3.3; SBN 414). Of course, not all ideas of potential pains and pleasures constitute such "prospects". While the *believed* idea of a pleasant drink in my fridge may produce a motivating desire to open the fridge, merely imagining a drink there would be unlikely to do so. It is generally assumed that, as Darwall suggests, Hume considers that only those ideas which "constitute beliefs" are potentially

motivating (“Motive and Obligation in Hume’s Ethics”, 423). Similarly, Stroud understands that to have a prospect in mind is to “believe something” (*Hume*, 157). Cohon thinks Hume allows that some non-believed ideas may *cause* passions, but still understands that only a believed idea can cause a passion which is “sufficiently strong to cause me to act” (*Hume’s Morality*, 42).

If this general view is correct, then Hume must show that we *believe* that performing formally just actions is either certain or likely to cause pleasure or prevent pain, if he is to explain why we are motivated to perform them. Hence, Garrett takes Hume to argue that we “come to believe” that it is in our interests to resolve to perform *all* just actions without considering the consequences (“The First Motive to Justice”, 274). By Stroud’s interpretation, Hume thinks we have the “false belief” that society will collapse if we omit *any* act of justice (*Hume*, 214).

We might assume that the relevant belief for the motivation of a formally just action is that its performance will produce the pleasure of approbation. However, this is problematic, for Hume believes that we *only* approve of those virtues which are “useful to society” – of which justice is his paradigm example – because they cause us to feel sympathetic pleasures with “the happiness of strangers” (T 3.3.6.2; SBN 619). *Prima facie*, we should expect that any future action that we believe will cause only unhappiness would also be one that we believe will produce not approbation but disapprobation. To understand why this is not the case, we must ask precisely what Hume means by a “prospect” of pain or pleasure, and we must look to the causes of the moral passions. We will see that not all “prospects” of pain and pleasure are beliefs, because Hume allows that some ideas cause motivating passions even where they are *not* believed. True, belief is “*almost* absolutely requisite to the exciting our passions” (T 1.3.10.4, SBN 120, my italics). However, a close examination of Hume’s theory of causal reasoning will demonstrate that some passions, including moral sentiments, may be

“excited” in the absence of belief.

Hume’s account of causal reasoning combines two distinct theses: that of the habitual association of ideas and that of the transference of vivacity between perceptions. The former is “a gentle force, which commonly prevails” between ideas whenever they are related by cause and effect (T 1.1.4.1; SBN 10). This relation of cause and effect is due to repeated experiences of witnessing one kind of object follow another kind, so that we come to habitually associate the two. Wherever two such object kinds are strongly associated by custom in this way, any perception of one will produce an idea of the other. For example, if I have frequently seen unsupported objects fall downwards, then any perception of an unsupported object will cause me to form an idea of it falling.

Importantly, however, this association of ideas is never sufficient to cause a *belief* that a token object is a cause or effect. Hume’s second thesis is that, even where an idea is causally associated with a present perception, it will only be believed if sufficient “vivacity” is transferred between the two perceptions (T 1.3.8.2; SBN 98). Here we need not consider precisely what Hume means by “vivacity”, beyond the fact that vivid ideas feel real to us in an approximation of the way that impressions feel real.⁸ Impressions, such as sense perceptions and feelings, are always vividly real. Hume’s thesis is that, where an idea is brought to mind because it is habitually associated with a “present impression”, some of the vivacity of that impression will generally be transferred to the idea, so that the idea also becomes vivid: it becomes a belief (T 1.3.7.5; SBN 96).

It is however possible for an impression to be associated with an idea, but without the idea gaining enough vivacity to constitute a belief. This may occur where *incompatible* ideas are associated with one’s current impressions. For example, Hume describes how someone “hung out from a high tower in a cage of iron cannot forbear trembling, when he surveys the precipice below him, tho’ he knows himself to be perfectly secure from falling” (T 1.3.13.10;

SBN 148). The idea of falling is associated with the impression of distance between oneself and the ground, just as the idea of solidity is associated with the impression of iron. Both ideas come to mind, but are incompatible as beliefs, because the solidity of the iron prevents falling. Therefore, only the idea of solidity is made vivid enough to be believed.

Hume claims that the idea of falling, although *not* vivid enough to be believed, has *some* “force and vivacity, which make it superior to the mere fictions of the fancy” (T 1.3.13.9, SBN 148). Part of this superiority, presumably, is that it can cause fear and “trembling” (T 1.3.13.10; SBN 148). This fear is not caused by a *general* belief, such as that suspension usually causes falling, but by the idea of the man himself falling. We might call this idea a *quasi*-belief, due to its moderate level of vivacity. It is clearly capable of causing passions. We will see that something similar can occur with the moral passions: approbation and disapprobation.

These passions are impressions, but not “original” impressions, like those of our senses. Instead, they are “secondary” impressions, which are “such as proceed from some of these original ones, either immediately or by the interposition of its idea” (T 2.1.1.1; SBN 275).

Hume classifies different secondary impressions according to their causes. He defines approbation and disapprobation as those pleasures and pains which are caused by actions and characters when we consider them “in general, without reference to our particular interest” (T 3.1.2.4; SBN 472). The moral passions are unselfish ones, and Hume ultimately explains them by reference to our sympathetic reactions to the pleasures and pains of others (T 3.3.6.2; SBN 619). However, here Hume claims they are caused by considering actions and characters not only *disinterestedly*, but also *generally*. He then says that “virtue 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; SBN 475). This does not suggest that moral sentiments

are caused by considering the *consequences* of token actions, sentiments or characters, whether to ourselves or to others. It suggests that we feel these passions simply when we consider what general *kinds* of actions, sentiments or characters we are witnessing. This interpretation is corroborated by Hume's ultimate explanation of approbation as the pleasure caused "by the mere survey" of any quality of mind which is "naturally fitted" to be useful or agreeable (T 3.3.1.30; SBN 591). By this point he has clearly indicated, as we saw in his discussion of "virtue in rags", that to be "fitted" to cause an outcome is to be of a kind which tends to cause that outcome, *ceteris paribus* (T 3.3.1.20; SBN 584).

Why should approbation be caused by any quality of mind which *tends* to be useful or agreeable? The answer is that such qualities frequently cause happiness to those around oneself, so that, wherever one takes an action or motive to be naturally "fitted" to be useful or agreeable – to be of a generally useful or agreeable kind – one will form an idea of the happiness of others by customary association. *Any* token perception of a useful or agreeable kind of action – such as a benevolent or just action – may produce an idea of such happiness. Here we can ask whether an *unbelieved* idea of the happiness of others can cause approbation, just as an unbelieved idea of falling can cause fear. Hume argues that it can.

He claims that we feel a "delicate sympathy" with someone *wherever* we witness an object which has "a tendency to produce pleasure in the possessor, or in other words, is the proper *cause* of pleasure" (T 3.3.1.8; SBN 577). One feels pleasure at the *tendency* of objects to directly please third parties, even where no such direct pleasure results. Hume later gives a concrete example: "A house, that is contriv'd with great judgment for all the commodities of life, pleases us upon that account; tho' perhaps we are sensible, that no-one will ever dwell in it" (T 3.3.1.20; SBN 584).

Hume argues that the "same principle produces, in many instances, our sentiments of morals, as well as those of beauty" (T 3.3.1.9; SBN 577). It is with this principle that Hume

explains “Virtue in rags”:

The imagination has a set of passions belonging to it, upon which our sentiments of beauty much depend. These passions are mov’d by degrees of liveliness and strength, which are inferior to *belief*, and independent of the real existence of their objects. Where a character is, in every respect, fitted to be beneficial to society, the imagination passes easily from the cause to the effect, without considering that there are still some circumstances wanting to render the cause a compleat one. *General rules* create a species of probability, which sometimes influences the judgment, and always the imagination. (T 3.3.1.20; SBN 585)

Here, Hume argues that any character of a generally useful kind will cause a passion of approbation, even where the character is prevented by circumstance from benefitting anyone. As a “general rule”, socially beneficial characters and motives – which certainly include *just* characters and motives – cause happiness to others. Therefore, *any* idea of a just motive, or of its associated action, will produce an idea of such happiness, with which we are pleased by a “delicate sympathy”. This in turn produces approbation.

Therefore, wherever we contemplate the idea of performing a formally just action, we will form an idea of happiness as caused by that action. This idea, being unbelieved, cannot directly motivate us to act. However, just like the idea of falling from the iron cage, it *can* indirectly motivate in such cases, by acting as a quasi-belief: an idea of insufficient vivacity to be believed but sufficient vivacity to produce passions. Ideas of the happiness of strangers, like those of falling, are highly emotive ones, and so it should be no surprise that such an idea can act as a quasi-belief in this way. This quasi-belief in the prospect of happiness arising

from one's action is sufficient to produce approbation towards the idea of so acting. Similarly, the thought of *not* performing the action will produce a quasi-belief in the prospect of the unhappiness of others, sufficient to produce disapprobation. These moral sentiments are directly motivating. This explains the motive to perform formally just actions. We perform them simply because we deem it moral to do so and immoral not to.

6. Conclusion

We should understand Hume to be speaking literally when he claims that “moral distinctions arise, in a great measure, from the *tendency* of qualities and characters to the interest of society, and... ’tis our concern for that interest, which makes us approve or disapprove of them (T 3.3.1.11; SBN 579, my italics). Once we have come to associate just actions with causing happiness, we approve of *every* motive to act justly and disapprove of *every* motive to act unjustly. Equally, because these motives are strongly associated with the actions themselves, we come to approve of *every* just action and disapprove of *every* unjust action. Therefore, we feel duty-bound to act justly even where we believe that the action holds no prospect of pleasure for ourselves or others.

Not only does this interpretation resolve the problem of formally just actions, but it resolves the original motive problem as well. This resolution rests on the fact that Hume only claims, as Garrett (“The First Motive to Justice”, 260) also notes, that we *naturally* have no “real or universal motive for observing the laws of equity, but the very equity and merit of that observance” (T 3.2.1.17; SBN 483). While Hume requires that we have a motive to be just which is distinct from moral duty, this motive may occur *artificially*. I have argued against Garrett that Hume does not suggest that artifice leads us to develop an entirely new *kind* of motive, whereby we desire to avoid assessing the consequences of just actions. Rather, he simply means that the rules of justice were artificially created, and were initially

followed from self-interest rather than by instinct. Once we are accustomed to follow the rules, we approve of the self-interested motivation to do so, in the same way that we approve of the instinctively arising motivations to help others or to act in one's children's interests.

This interpretation explains our approval of *all* virtuous actions without requiring any addition or amendment to Hume's theory to explain justice. It explains why Hume treats the artificial virtues before the natural virtues, and it is faithful to his text. However, I must allow that there is still more to be explained. I have not addressed Hume's discussion of the "common point of view", which he believes us to take when we are considering matters of morality (T 3.3.1.30; SBN 591). This is an important aspect of Hume's theory, and a full account of his theory of approval would address this in detail. Equally, there are important implications of this to Hume's broader accounts of belief, motivation and the passions, which require further exploration. Nevertheless, I believe that this interpretation of Hume's theory of approval and moral motivation is a promising one, and one which warrants further examination.

NOTES

I would like to thank Harold Noonan and Neil Sinclair, my PhD supervisors at the time of writing, along with many others in the University of Nottingham's Department of Philosophy for their helpful comments. I am also grateful for the support and guidance of the editors and referees at *Hume Studies*.

¹ References to the *Treatise* are to Hume, *A Treatise of Human Nature*, ed. Norton and Norton, hereafter cited in the text as "T" followed by Book, part, section, and paragraph numbers, and to *A Treatise of Human Nature*, ed. Selby-Bigge, revised by Nidditch, hereafter cited in the text as "SBN" followed by page numbers. References to the second *Enquiry* are to Hume, *An Enquiry concerning the Principles of Morals*, ed. Beauchamp, hereafter cited in the text as "EPM" followed by section and paragraph number, and to Hume, *Enquiries concerning the Human Understanding and concerning the Principles of Morals*, ed. Selby-Bigge, revised by Nidditch, hereafter cited in the text as "SBN" followed by page numbers.

² My thanks to an anonymous reviewer for stressing the importance of this fact to Hume's account of justice.

³ This is suggested by Krause, "Hume and the (False) Luster of Justice"

⁴ This possibility has been suggested, within varying broader interpretations, by at least the following: Baier, *A Progress of Sentiments*; Besser-Jones, "The Role of Justice in Hume's Theory of Psychological Development"; Bricke, *Mind and Morality*; Cohon, *Hume's Morality*; Darwall, "Motive and Obligation in Hume's Ethics"; Garrett, *Hume*; Stroud, *Hume*.

⁵ My thanks to two anonymous reviewers for pressing this point.

⁶ See T 3.3.1.9, 3.3.1.10, 3.3.1.11, 3.3.1.12, 3.3.1.13, 3.3.1.14, 3.3.1.19, 3.3.1.23, 3.3.1.25, 3.3.1.27, 3.3.1.28, 3.3.2.15, 3.3.3.3, 3.3.4.5, 3.3.4.11, 3.3.5.1, 3.3.6.1, 3.3.6.4; SBN 577, 578, 579, 580, 580, 584, 586, 588, 589, 589, 601, 604, 610, 612, 614, 618, 620

⁷ References to *Of the Original Contract* are to Hume, “Of the Original Contract”, in *Essays, Moral, Political and Literary*, 465-87.

⁸ Owen provides a helpful account of this feature of Hume’s theory of reasoning in *Hume’s Reason*, particularly Chapter 7

WORKS CITED

- Árdal, Páll S. *Passion and Value in Hume's Treatise*. Edinburgh: Edinburgh University Press, 1966.
- Ayer, A. J. *Hume*. Oxford: Oxford University Press, 1980.
- Baier, Annette. *A Progress of Sentiments*. Cambridge: Harvard University Press, 1991
- Baron, Marcia. "Hume's Noble Lie", in *Hume: Moral and Political Philosophy*, edited by Rachel Cohon, 273–90. Aldershot/Burlington: Ashgate/Dartmouth, 2001.
- Besser-Jones, Lorraine. "The Role of Justice in Hume's Theory of Psychological Development", *Hume Studies* 32 (2006): 153–276.
- Bricke, John. *Mind and Morality: An Examination of Hume's Moral Psychology*. Oxford: Oxford University Press, 2000.
- Cohon, Rachel. *Hume's Morality: Feeling and Fabrication*. Oxford: Oxford University Press, 2008.
- Darwall, Stephen. "Motive and Obligation in Hume's Ethics", *Noûs* 27 (1993): 415–48.
- Garrett, Don. *Hume*. Oxon: Routledge, 2015.
- Garrett, Don. "The First Motive to Justice: Hume's Circle Argument Squared", *Hume Studies* 33 (2007): 257–88.
- Gauthier, David. "Artificial Virtues and the Sensible Knave", *Hume Studies* 18 (1992): 401–28.
- Harris, James A. "Hume on the Moral Obligation to Justice", *Hume Studies* 36 (2010): 25–50.
- Harrison, Jonathan. *Hume's Theory of Justice*. Oxford: Clarendon Press, 1981.
- Hume, David. *An Enquiry concerning the Principles of Morals*. Edited by T. L. Beauchamp. Oxford: Oxford University Press, 2006

- Hume, David. *An Enquiry concerning the Human Understanding and concerning the Principles of Morals*. Edited by L. A. Selby-Bigge, revised by P. H. Nidditch. Oxford: Clarendon Press, 1975.
- Hume, David. "Of the Original Contract", in *Essays, Moral, Political and Literary*, revised edition. Edited by Eugene F. Miller, 465-87. Indianapolis: Liberty Classics, 1987.
- Hume, David. *A Treatise of Human Nature*. Edited by David F. Norton and Mary J. Norton. Oxford: Oxford University Press, 2000.
- Hume, David. *A Treatise of Human Nature*, 2nd ed. Edited by L. A. Selby-Bigge, revised by P. H. Nidditch. Oxford: Clarendon Press, 1978.
- Krause, Sharon R. "Hume and the (False) Luster of Justice", *Political Theory* 32 (2004): 628–55.
- Owen, David. *Hume's Reason*. Oxford: Oxford University Press, 2002.
- Stroud, Barry. *Hume*. Oxon: Routledge, 1977.
- Taylor, Jacqueline. "Justice and the Foundations of Social Morality in Hume's Treatise", *Hume Studies* 24 (1998): 5–30.