

HUME'S JUSTICE AND THE PROBLEM OF THE MISSING MOTIVE

IAN CRUISE

University of Alabama at Birmingham

The task that Hume explicitly sets himself in 3.2 of the *Treatise* is to identify the motive that renders just actions virtuous and constitutes justice as a virtue. But surprisingly, he never provides a clear account of what this motive is. This is the problem of the missing motive. The goal of this paper is to explain this problem and offer a novel solution. To set up my solution, I analyze a recent proposal from Geoffrey Sayre-McCord and illustrate what it gets right and what it gets wrong. I develop a solution that retains the benefits of his proposal while addressing its defects. The result is a significant advancement in our understanding of Hume's theory of justice.

HUME believes, plausibly, that each virtue has some characteristic motive. For example, we identify kind people as kind according to what motivates their characteristic behavior. While identifying the motive characteristic of certain virtues is relatively straightforward (the kind person, for example, is motivated by the needs of others), the question is more complicated when we consider what Hume calls the artificial virtues. What motivates the just person? In 3.2 of the *Treatise*, Hume seeks to identify the motive that constitutes justice as a virtue. But surprisingly, he never goes on to offer a clear account of what this motive is. This is "the problem of the missing motive."

Hume scholars have taken it upon themselves to propose accounts of what this motive must be to be consistent with Hume's other commitments. This is a serious challenge, however, because the problem of the missing motive intersects with so many other features of Hume's philosophy, many of which are themselves contested among Hume's commentators, including his account of the distinction between the artificial and the natural virtues, his account of the conventions of justice, and his motivational psychology.

Contact: Ian Cruise <ian.cruise44@gmail.com>

My goal in this paper is to offer a clear statement of the problem and a novel solution. Along the way, I will explain some of the advantages that my solution has over others that have been proposed in the literature, in particular the most recent developed solution given by Geoffrey Sayre-McCord. While the paper will focus on the problem as Hume develops it, the philosophical question with which Hume was engaged—the question of the motive characteristic of the just person—is of general interest to philosophers engaged with questions about the nature of virtue and the virtues.

Finally, before getting to Hume, I should set expectations in two ways. First, to understand the problem of the missing motive fully, the reader needs to be thoroughly acquainted with the argument of 3.2.1 of the *Treatise*. It is primarily in this section that Hume sets out the key desiderata for any solution to the problem. Because I intend this paper to be of interest not only to Hume scholars but also to philosophers interested in the virtue of justice, I work through (my reading of) this argument in some detail. That said, I also read the argument somewhat differently from some other Hume scholars, so it is important for me to articulate my interpretation since it informs my solution to the problem. As a result, my solution enters the scene only later on in the paper.

Second, as with nearly all interpretations of philosophical texts, I don't actually think that the text dictates a definite solution to the problem of the missing motive. In developing my own interpretation in this paper, I am not claiming that the Hume scholars who have proposed different solutions to this problem so far have simply misread or misunderstood Hume's texts. All of the proposals in the literature sit comfortably with some aspects of the text and uncomfortably with others. I will draw attention to some of the ways in which I think that my proposal does better than the extant options in the literature. But I will not claim definitively to refute any of the alternatives. Instead, and because I don't think that the text definitively supports any particular interpretation, I help myself to interpretive resources beyond the text in developing my proposal. In particular, I appeal to the plausibility of the resulting philosophical view in comparison with those that other interpreters attribute to Hume. The interpretation must, of course, be grounded in the text, but where the text is unclear, inconsistent, or simply silent, I take the task of the interpreter to be to *develop* the philosophical view being interpreted. With that in mind, let's turn to Hume.

1. The Task of *Treatise* 3.2

T 3.2, in which Hume develops his theory of justice, must be read in light of the problem that Hume sets himself in T 3.2.1. The task of T 3.2.1 is to argue that jus-

tice (if it is a virtue at all) must be an artificial virtue. It is worth recounting this argument because it sets up the quest for the motive of justice.

Hume begins by arguing that when we judge whether actions are praiseworthy or blameworthy, we attend not (in the first place) to the actions themselves but rather to the motives from which they are done (T 3.2.1.2 / SBN 477).¹ Hume's thought is the relatively common one that actions done from, say, a selfish motive don't merit praise. If I save someone from drowning only because I hope to get a reward, then my action is not praiseworthy. The action is virtuous, then, only if it is performed from a virtuous motive.

Hume concludes from this claim that "all virtuous actions derive their merit only from virtuous motives" (T 3.2.1.4 / SBN 478). We denominate actions as virtuous, or praiseworthy, when we take them to be motivated by virtuous motives. This raises the question of what these virtuous motives are that render virtuous actions virtuous. In particular, Hume wonders what motive could *originally* render actions of the relevant kind virtuous. Hume's concern is to determine which motives make virtuous actions virtuous *in the first place*.

Hume quickly dismisses a tempting thought: The motive that renders a virtuous action virtuous is "a regard to the virtue of [the] action" (T 3.2.1.4 / SBN 478). The person with this motive attends to the moral properties of her actions. Perhaps the person performs the action because it is required (the motive of duty), or because it is a good thing to do (the motive of value), or because it's the kind of thing a virtuous person would do (the motive of virtue), in each case read *de dicto*.² Hume thinks that this can't be the motive that *originally* renders virtuous actions virtuous. Here's how I understand the argument, which has come to be known as "the circle argument."

An action is originally made virtuous by its being done from a certain motive. But in order to perform an action from a regard to its virtue, it must already be virtuous. It isn't possible to perform an action from a regard to its virtue unless it is already virtuous because Hume is thinking of this "regard" as

1. This paper adopts the convention, standard within Hume scholarship, of citing Hume's *A Treatise of Human Nature* as "T," his *Enquiry Concerning the Principles of Morals* as "EPM," and his *A Dissertation on the Passions* as "DP." The numbers in citations to the *Treatise* refer to Book, part, section, and paragraph. "SBN #" refers to the page number(s) of the paragraph in question in the Selby-Bigge/Nidditch edition of the *Treatise*. References to the *Enquiry* are similar, though they only contain section and paragraph numbers. The one reference to the *Dissertation* contains only section and paragraph numbers and no Selby-Bigge/Nidditch citation since there is no Selby-Bigge/Nidditch edition of the *Dissertation*. I do, however, include the page reference for the Clarendon edition.

2. Hume later glosses this motive as "the sense of duty" (T 3.2.5.6; SBN 518–519). I think that Hume's argument applies more widely, though I can see why Hume would focus on duty given that his focus is on justice.

factive. That is, it follows from one's doing something from a regard to X that X (just as it follows from the fact that one knows that P that P). Hume writes, "Before we can have such a regard [i.e., the regard to the virtue of an action], the action must be really virtuous" (T 3.2.1.4 / SBN 478). It wouldn't be possible, on Hume's view, to perform an action from a regard to its virtue unless it were virtuous. Now, to say that the action must already be virtuous is to say that it must be an action that would be virtuous if someone were to perform it from a certain virtuous motive. Thus, to perform an action from a regard to its virtue would be to perform it from a regard to the fact that it would be virtuous if someone were to perform it from a certain motive. This motive, then, has another motive as part of its content, and it is that motive, whatever it is, that renders the action virtuous. Call this motive "the embedded motive." The problem with claiming that this motive (the motive of a regard to the virtue of the action) is the motive that originally renders an action virtuous is that the embedded motive would have to be the regard to the virtue of the action. But then the embedded motive would have a further embedded motive in it, which would also have to be a regard to the virtue of the action. And so on.

Notice that this way of understanding the argument makes clear why proposing a different formulation of this motive won't help. One might, for instance, think that this motive isn't a regard to the virtue of an action but instead merely the belief that an action is virtuous. A belief that the action is virtuous isn't factive. But we get the same regress problem on either formulation. Here's why. Suppose that someone performs an action because she believes it to be virtuous. As Hume understands it, this would mean that her motive for performing the action is that she believes that it would be virtuous if someone were to perform it from a certain motive. Call this the modified motive. Yet again, we have an embedded motive. If the embedded motive were the modified motive, then her motive would be that she believes that it would be virtuous if someone were to perform the action from the belief that it would be virtuous if someone were to perform the action from a certain motive. This is quite a complicated motive. And yet again, we have a more deeply embedded motive. If this motive were the modified motive, then the motive would be more complicated still. The problem, of course, isn't that the motive is complicated. Rather, it's that building in more and more deeply embedded motives never illuminates what the motive that renders the action virtuous is. And the reason is that if the embedded motive (at any level of depth) is the modified motive, we will always have a more deeply embedded motive in need of explication.

Hume claims that to think that a regard to the virtue of the action is the motive we're after would be "to reason in a circle" (T 3.2.1.4 / SBN 478), which is why this argument is known as the circle argument. As should be clear from the

way I outline the argument, I think the problem is better thought of as a regress than a circle. But we can think of the problem as a circle as follows. Suppose we claim that a regard to the virtue of the action is the motive that originally renders an action virtuous. But in order even to make sense of what this motive is, we need to get clear on another motive, namely, the embedded motive. But if the embedded motive is just a regard to the virtue of the action, then we haven't gotten anywhere. We've simply gone in a circle.³ I think that this is probably how Hume was thinking of the problem. That said, my view is that the explanation for why we haven't gotten anywhere is the regress of more and more deeply embedded motives, which never illuminates what the motive is that originally renders an action virtuous.

A regard to the virtue of the action, then, can't be the motive that originally renders *any* virtuous action virtuous. Hume writes, "*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, original emphasis). Importantly, it doesn't follow that the regard to the virtue of the action isn't a virtuous motive. It just means that it can't be the one that *originally* makes virtuous actions virtuous. It might very well be virtuous to act on this motive. But this motive can't be the one that makes, for example, just actions virtuous.

On that note, let's follow Hume in turning to the virtue of justice more specifically. As the previous discussion indicates, the virtue of justice is constituted by a disposition to act on whatever motive it is that renders just actions virtuous. For now, let's call this motive, whatever it is, the motive of justice. What is the motive of justice? As Hume writes, "'Tis requisite . . . to find some motive to acts of justice and honesty, distinct from our regard to the honesty; and in this lies the great difficulty'" (T 3.2.1.10 / SBN 480). Hume approaches this challenge, first, by asking whether the motive of justice is constitutive of a natural virtue or an artificial virtue.

There's some degree of scholarly disagreement about how precisely to distinguish between the natural virtues and the artificial virtues, but the basic idea is the following.⁴ Artificial virtues, Hume says, "produce pleasure and approbation by means of an artifice or contrivance" (T 3.2.1.1 / SBN 477). The "artifice or contrivance" in question, we learn later, is a cooperative social convention. Artificial virtues can be contrasted with natural virtues, then, in that the natural

3. Cohon (2008) and Sayre-McCord (2016) outline the argument along these lines.

4. J.L. Mackie (1980), for instance, thinks that there is, after all, no distinction between the two. Others who have weighed in on this distinction include Fieser (1997), Cohon (2006), Abramson (2015), and Kelahan (2018).

virtues secure approval (“produce pleasure and approbation”)⁵ independently of the operation of a convention, though conventions can affect the *degree* of approval and the content of the virtue in question.⁶ Our approval, as such, of the natural virtues is not dependent on their being connected with or founded on any convention.

Moreover, the value of the artificial virtues is a product of consistently abiding by the demands of the relevant conventions. Particular acts in accordance with the demands of justice might be good for no one. But nonetheless the public and each individual benefits from each of us consistently abiding by the demands of justice. Hume writes, “But 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 / SBN 497–498). The contrast with the natural virtues here is *not* that particular acts from a natural virtue always succeed in securing the good at which they aim. Rather, the contrast is that the natural virtues always *aim at* the good in question with each particular act (even if not every particular act actually ends up successfully securing the good). The same is not true of the artificial virtues. The artificial virtues aim at the good overall through the operation of the conventions, but they aim at adherence to the relevant conventions with each particular act. The good is a product of the successful operation of the convention that requires each particular act. The missing motive, then, must involve consistent adherence to the demands of the relevant conventions.

Finally, and this is a bit more controversial (in that other commentators don’t typically use this point to mark the distinction), we contrast artificial and natural virtues according to the motives that constitute them. In particular, artificial virtues are constituted by artificial motives, and natural virtues are constituted by natural motives. An artificial motive is one that is available only in the presence of conventions (in the sense that the conventions make the motive possible). A natural motive is available independently of conventions.

This distinction between the artificial and natural virtues might be thought controversial because of the following passage:

The *only* difference betwixt the natural virtues and justice lies in this, that the good, which results from the former, arises from every single

5. The relevant approval is approval from what Hume calls the general or common point of view ([T 3.3.1.15 / SBN 581–2] and [EPM 9.6 / SBN 272–3]). The adoption of the general point of view is meant to correct for some of the biases inherent in sympathy. Any appeal to “approval” or “approbation” from here on out should be understood as “approval from the general point of view” or “approbation from the general point of view.”

6. Kate Abramson makes this last point in (2015: 352).

act, and is the object of some natural passion: Whereas a single act of justice, consider'd in itself, may often be contrary to the public good; and 'tis only the concurrence of mankind, in a general scheme or system of action, which is advantageous. (T 3.3.1.12 / SBN 579–580, my emphasis)

If this were the *only* difference between the natural and artificial virtues, then that seems to leave no room for distinguishing them on the basis of the motives that constitute them.⁷

I think there is reason to think that Hume is mistaken about his own view (or what his view should be) in this passage. To see why, we need to turn to the next part of the argument of T 3.2.1. Hume's next move is to argue that justice, if it is a virtue at all, must be an artificial virtue. He argues, by elimination, that none of three plausible natural motives could constitute the virtue of justice.

What are the natural motives that he considers and why does he think that none of them is the motive that renders just actions virtuous? Hume considers three different natural motives: self-love, "a regard to publick interest" (T 3.2.1.11 / SBN 480) (which I'll call "public benevolence"), and "a regard to the interests of the party concern'd" (T 3.2.1.13 / SBN 482) (which I'll call "private benevolence"). He then notices that it is possible for a person to be acting from each of these motives and yet fail to act justly or to be acting justly and yet not be acting from one of these motives.

First, consider self-love. According to Hume, self-love, at least "when it acts at its liberty . . . , is the source of all injustice and violence" (T 3.2.1.10 / SBN 480). It is therefore possible to be acting from self-love and yet not be acting justly. It is easy enough to come up with cases to illustrate the point. One might, for example, steal someone else's property from self-love.⁸

7. To be clear, the reference to "some natural passion" doesn't support the view that artificial and natural virtues can be distinguished on the basis of the former being constituted by artificial motives and the latter natural motives. Hume's claim is that "the good, which results from [the natural virtues] . . . is the object of some natural passion." This simply means that our approval of the good that the natural virtues produce is not convention-dependent. The reference to a "natural passion" here is not a reference to a natural motive.

8. There are a number of complexities lying in the background of Hume's brief dismissal of self-love as the motive of justice here. The view that justice is constituted by self-love is perhaps most strongly associated with Hobbes, though the view goes back to others including Epicurus and the Glaucon of Book II of the *Republic*. Given that self-love ultimately does play a role in Hume's genealogy of justice (as the motive for the establishment of the many conventions discussed in T 3.2), many have thought that ultimately Hume accepts that some modification of self-love is the motive of justice. This interpretation is perhaps most strongly associated with David Gauthier (1979; 1992) and Annette Baier (1991; 2010). Thus, the reader should be aware that Hume's quick elimination of self-love as a contender for the motive of justice should not be read to imply that Hume doesn't think that more needs to be said about the relationship between justice and self-love.

Second, consider public benevolence. The main reason that this can't be the motive is that most interactions in which justice is at stake (the repayment of a loan, for instance) simply don't engage the interest of the public at all, and so there's no room for benevolent concern for the public. We all know that, and so one might be acting justly and from a virtuous motive (in repaying the loan) and yet not be acting from a concern to benefit the public.⁹

Third, consider private benevolence. Hume provides several examples to illustrate why a person acting from private benevolence might be motivated not to do as justice requires. He writes:

For what if he be my enemy, and has given me just cause to hate him? What if he be a vicious man, and deserves the hatred of all mankind? What if he be a miser, and can make no use of what I wou'd deprive him of? What if he be a profligate debauchee, and wou'd rather receive harm than benefit from large possessions? What if I be in necessity, and have urgent motives to acquire something to my family? In all these cases, the original motive to justice wou'd fail; and consequently the justice itself, and along with it all property, right, and obligation. (T 3.2.1.13 / SBN 482)

I have no private benevolence for my enemy, and yet I might repay a loan to him from the motive of justice. My private benevolence for a "profligate debauchee" might rather lead me to take some of his possessions, since he apparently makes himself worse off by having them, and yet I would refrain from his possessions from the motive of justice. Thus, there are cases in which one could act from the motive of justice and yet not be acting from private benevolence.

The conclusion of this argument is that justice, if it is a virtue at all, must be an artificial virtue. None of the plausible natural motives constitutes the virtue of justice.

The reason that we need the distinction between artificial and natural virtues on the basis of their being constituted by artificial and natural motives is that otherwise we couldn't make sense of this argument. It is a crucial feature of this argument that the motives Hume considers are *natural motives*, potentially constitutive of natural virtues, which might then render justice a virtue. In dismissing those motives, Hume is able to rebut the thought that justice is a natural virtue. As a result, justice must be an artificial virtue, constituted by an artificial motive. If an artificial virtue could be constituted by a natural motive that simply operates in the presence of a convention, then Hume's dismissal of the three

9. What I've said here about public benevolence will suffice for now to understand the thrust of Hume's argument from elimination, but given that the motive I will ultimately defend is related to the regard to the interest of the public, I will return to more of Hume's comments about this motive later on.

natural motives (self-love, public benevolence, and private benevolence) would be much too quick. And in any case, Hume dismisses these motives even as they are operative within clearly convention-dependent contexts, such as the case of repaying a loan (T 3.2.1.11 / SBN 480–481). Without this distinction between the artificial and natural virtues available, the argument from elimination in T 3.2.1 doesn't work.

Because no plausible natural motive constitutes the virtue of justice, according to Hume, justice must be an artificial virtue. He writes, "Unless ... we will allow, that nature has establish'd a sophistry, and render'd it necessary and unavoidable, we must allow, that the sense of justice and injustice is not deriv'd from nature, but arises artificially, tho' necessarily from education, and human conventions" (T 3.2.1.17 / SBN 483). But given 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), there must be some motive that renders just actions virtuous and, more to the point, that constitutes justice as a virtue. The problem of the missing motive is that Hume never clearly tells us what this motive is, despite his claim that "'Tis requisite...to find some motive to acts of justice and honesty, distinct from our regard to the honesty" (3.2.1.10 / SBN 480).

But one might think that there's still a puzzle here. If the motive must be "*in human nature*," then one might be led to think that virtues could only come from natural passions, which would leave us with only the option of a natural motive operating in the presence of a convention. I've argued, however, that Hume has to dismiss that option in order to make the argument of T 3.2.1 work. So how could the motive of justice be both in human nature and yet not a natural motive? The answer, I believe, and this seems the route that some commentators go, is to interpret the motive as a *modification* of a natural motive. Thus, we have some motive in human nature, and it gets modified in some way by education and convention. The new motive, the modification of the natural passion, is an artificial motive. With that in mind, let's turn to a recent solution.

2. Sayre-McCord's Solution

Geoffrey Sayre-McCord (2016) is the most recent major contributor to this debate.¹⁰ On his view, for Hume, what the just person is motivated to *do* is to

10. Other important contributors include Baron (1982), Cohon (2008), Postema (1988), Baier (1991; 2010), Gauthier (1979; 1992), Darwall (1993; 1995), Garrett (2007), and Harris (2010). Baron and Cohon think that the missing motive is the motive of duty. This proposal looks to be at least initially in tension with the circle argument. The circle argument seems to rule out the motive of duty as the motive that originally renders just actions virtuous. Cohon, to her credit,

do one's share in the maintenance of the conventions that generate the demands of justice by following their rules. And the just person's *motive* is a sense of reciprocity. Sayre-McCord writes:

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, self-interest, or a fetish for the rules) and those who are, in conforming, being just (i.e., 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 they will not be manifesting the virtue of justice. (2016: 458)

There are a few points worth emphasizing in this passage. First, the just person isn't moved to abide by the rules of just any convention. Conventions can be

has a response to this concern (that the circle argument only rules out the motive of duty as the motive constitutive of the natural virtues), but this response still requires reading Hume in an unnatural way.

Postema, Baier, and Gauthier think that the motive is a form of redirected self-interest. Each person's self-interest is redirected by effective conventions that channel individual self-interest in socially useful ways. One might think of this as a kind of invisible hand view. The major issue for this view is that self-interest, whether redirected or not, seems, if it is a virtue at all, to constitute a virtue like prudence, not justice. Of course, Hume might simply have an implausible view, but if we can find a more plausible option consistent with the text, then we should, on grounds of charity, attribute it to Hume.

Darwall and Garrett think that the motive is a commitment to abiding by the rules of the conventions of justice, a motivation that Darwall calls "the motivational state of rule-regulation" (1993: 437). Darwall himself raises the main objection to this view: It isn't clear how Hume could accept this as a motive to any action given his motivational psychology. Because, for Hume, any action must be motivated by prospective pleasure or the avoidance of prospective pain (T 2.3.9.7–8 / SBN 439), it isn't clear how a commitment to abiding by the rules that looks no further than the rules themselves could serve as a motive for Hume. That said, Garrett disputes the reading of Hume according to which all action must be motivated by prospective pleasure and pain, and he thinks that regardless, the motivational state of rule regulation is consistent with Hume's motivational psychology. More on Darwall and Garrett in the main text shortly.

Harris argues that, when he comes to justice, Hume abandons his view that all virtues have characteristic motives. This revisionary interpretation entails that there is, after all, no problem of the missing motive, which would require reading Hume as setting up a problem that he ultimately doesn't think requires a solution.

I don't intend these criticisms to be decisive. But rather than spend the whole paper carefully objecting to each of the major solutions to the problem of the missing motive on offer, I simply note that each of them faces important problems, problems that, I believe, my own solution doesn't face.

good or bad, and the just person, presumably, is not motivated to abide by morally objectionable conventions (e.g., slavery conventions). Such a person would not be exhibiting the virtue of justice but rather a kind of rule worship. Sayre-McCord captures this point by insisting that, for Hume, the conventions in question must be mutually advantageous.¹¹

Second, the just person's motive is conditional on others abiding by the conventions as well. Just people need have no commitment to abiding by the rules when others are not abiding by the rules, especially in the face of being consistently stolen from, lied to, and otherwise made to sacrifice their immediate interests when others are not similarly willing. This mindless adherence to the rules is no virtue.¹²

Because, for Sayre-McCord, the just person conditions her willingness to do her share in the maintenance of the conventions on others doing their share and because it is for this reason that she does her share, her motive is one of reciprocity. She values reciprocating the work of others in the joint project of maintaining the conventions and acts accordingly.

Sayre-McCord's proposal has much to recommend it. He identifies a clearly artificial motive (since a concern to reciprocate the work of others in the maintenance of a convention depends on there being a convention), which is at least arguably constitutive of an artificial virtue. This motive would universally motivate actions in accordance with the demands of justice, which saves it from the kind of criticism Hume levels against the motives he dismisses with the argument from elimination in T 3.2.1. It is also a non-moral motive, in the sense that it doesn't have distinctively moral concepts as part of its content, which saves it from the clutches of the circle argument.

11. This is a joint in the view, both as a matter of interpretation and substantively. Not all philosophers think that Hume's theory of justice makes the normative standard of justice one of mutual advantage. See, for example, Salter (2012). Others think that Hume is a mutual advantage theorist and criticize him for that reason. Mutual advantage theories of justice are open to, for example, the objection that they can't explain why those who can't or don't contribute to the productive process (such as those with certain severe disabilities and people outside our own societies, respectively) are excluded from the scope of justice. See, for example, Barry (1989: 162–163). It is open to those who work on Hume's theory of justice, both its interpretation and its development, to defend a different normative standard that the conventions must meet.

12. Sayre-McCord criticizes both Darwall and Garrett for failing to accommodate these points in their views. See Sayre-McCord (2016: 457; 466, fn. 57) for the attribution and for the criticism. Luckily, I don't think that either Darwall or Garrett is committed to the simple form of this view. Darwall (1993: 437) isn't committed to the view because he allows that more needs to be said to fill in the precise content of the motive, including that the commitment to the rules is conditional and that the rules themselves must meet some moral standard. Garrett thinks that, for Hume, the motive must result from the recognition of "the obligating authority" of the rules (2007: 274). One could charitably interpret Garrett as thinking that the obligating authority of the rules of justice depends on their meeting certain conditions, including that people generally adhere to them and that they meet certain normative standards. If we interpret him in this way, Garrett can similarly make the motive conditional in a way that avoids the objection.

While Sayre-McCord's proposal is arguably the most independently attractive proposal for what the motive of justice actually is in this literature, it faces two main problems. First, there's just not much good textual evidence for it. It is debatably an attractive proposal for the motive of justice, but it isn't obviously an attractive proposal for Hume's motive of justice. Of course, the fact that Hume's missing motive is *missing* gives the interpreter some leeway to go beyond the text in developing an account, but the motive should, ideally, have some stronger grounding in the text. Hume simply doesn't talk about reciprocity in his discussion of justice, except at best tangentially when he notes, in various places, that the rules of justice are dependent on conventions, which themselves depend for their stability on each of us acting in accordance with them (e.g., [T 3.2.2.10 / SBN 490] and [EPM App. 3.5 / SBN 305]). But he nowhere says that the just person acts *because* she wishes to reciprocate the work of others.

Second, the proposal falls victim to a problem that Darwall identifies. Darwall thinks that, for Hume, the motive of justice must be the motive to regulate one's actions by rules with no further end in view, but he thinks that this motive is not available to Hume given his motivational psychology. According to Hume, except in the case of those passions that come "from a natural impulse or instinct" (which Hume thinks are "unaccountable" [T 2.3.9.8 / SBN 439]), action must be motivated by some prospective pleasure or the avoidance of some prospective pain (T 2.3.9.7–8 / SBN 439). With regard to Sayre-McCord's proposal, it isn't clear how a concern for reciprocity, in and of itself, aims at some prospective pleasure or the avoidance of some prospective pain. Darwall is content to conclude that Hume's view is simply inconsistent. But the correct account of the missing motive would, ideally, be consistent with Hume's motivational psychology. Given that a concern for reciprocity, no less than the motive to regulate one's conduct by the rules of the conventions, doesn't obviously aim at any pleasure or the avoidance of any pain, Sayre-McCord's account doesn't look like the best one *to attribute to Hume*.

So then altogether, what we're after is a non-moral, universal, artificial motive to acts of justice that constitutes a virtue and at once finds some plausible grounding in the text and harmonizes with Hume's motivational psychology. Is such a motive available?

3. A New Proposal

My proposal is that the missing motive is a kind of redirected concern for the public interest. The redirection is important because the motive obviously can't be the natural version of the motive (i.e., public benevolence) without being inconsistent with the fact that the motive can't be any of the ones that Hume dis-

misses in T 3.2.1. But a redirected concern for the public interest does not violate this desideratum.

Let me sharpen this motive up a bit. First, the concern for the public interest is redirected in the sense that the conventions give us a longer-term perspective on the public interest. We need not respect the property of others, keep our promises, etc., because we expect that each particular act directly promotes the public interest (T 3.2.2.22 / SBN 497–498). Rather, we recognize that consistently adhering to the demands of the conventions is in the long-term public interest.¹³ It is this effect of the conventions on our motive that renders this motive artificial and so a candidate for the motive constitutive of the artificial virtue that we're after.

Second, the motive is conditional. Our concern to promote the public interest by abiding by the rules of the conventions is conditional both on others also abiding by the conventions and on the content of the rules. Our concern to promote the public interest does not entail that we're willing to abide by the rules when others are not similarly willing (though, crucially, we don't adhere to the rules *for the reason that* others do, as in Sayre-McCord's proposal). And if the rules of the conventions did not themselves secure our approbation, then we would not think it virtuous to abide by them. There's no virtue in abiding by morally objectionable rules.

What exactly is this "public interest" that the person with the virtue of justice is supposed to be concerned to promote? On my reading, a concern with the public interest, for Hume, is not just a concern for the interest of the public understood in aggregate, as, for example, the utilitarian must understand it.¹⁴ As I've previously noted, Hume writes, "But 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 / SBN 497–498).¹⁵ A concern for public interest, Hume thinks, involves a concern both for society as a

13. A.D. Woozley proposes that the motive is "a sense of common interest (which is a recognition of the mutuality of one's own and others' interests)" (1978: 87). This proposal is ambiguous between mine and Sayre-McCord's, but regardless, Woozley objects to it on the grounds that not all acts in accordance with the demands of justice serve the common interest. The point I'm making here shows why this objection is off the mark. The motive isn't sensitive to whether each and every act is in the public interest. It is sensitive instead to whether consistent adherence to the conventions is in the public interest.

14. Sayre-McCord (1995) offers a sweeping critique of the utilitarian interpretation of Hume's moral philosophy, which I largely agree with.

15. Hume also makes the point that justice is good not just for the society as a whole but also for each individual at (T 3.2.2.20 / SBN 496) and (T 3.1.1.12 / SBN 579–80). Moreover, later in the T 3.2.2.22 passage, Hume writes that "even every individual person must find himself a gainer [from justice], on ballancing the account" and that justice "is infinitely advantageous to the whole, and to every part." The point that justice is good for each individual seems to have been important to Hume.

whole and for each individual. This concern, then, does not allow us to sacrifice the interests of some just to increase the aggregate well-being, at least not below some minimum threshold (which Hume does not clearly define). Hume here prefigures, to some extent, Rawls's concern that utilitarianism fails to respect "the distinction between persons" (1999: 24).

We should read a concern for the public interest as a concern for the well-being of each person whose interests are affected by the operation of a convention. The motive, then, fully spelled out, is a conditional concern to abide by the rules of conventions to which one is subject that promote the well-being of each person whose interests these conventions affect. Abiding by the rules is what the person with the virtue of justice is motivated *to do*, and a conditional concern for the well-being of each person affected by the conventions is *what motivates* abiding by the rules.

My proposal has the same strengths as Sayre-McCord's. It is not one of the motives that Hume dismisses in T 3.2.1 because a *suitably conditional and redirected* concern for the public interest is not one of the motives Hume dismisses. It universally motivates action in accordance with the demands of justice since the demands of justice are fixed by the demands of the relevant conventions, and what the motive motivates us to do is to abide by these conventions. It is non-moral because it makes no reference to the moral status of the actions themselves. And it is an artificial motive the disposition to act on which plausibly secures moral approbation, thus rendering the disposition to act on the motive an artificial virtue. The main questions we must ask of my proposal are 1) whether it can find some plausible grounding in the text and 2) whether it can be squared with Hume's motivational psychology.

Regarding the first, a concern with my view is that it seems in tension with a crucial passage. Hume writes:

experience sufficiently proves, that men, in the ordinary conduct of life, look not so far as the public interest, when they pay their creditors, perform their promises, and abstain from theft, and robbery, and injustice of every kind. That is a motive too remote and too sublime to affect the generality of mankind, and operate with any force in actions so contrary to private interest as are frequently those of justice and common honesty. (T 3.2.1.11 / SBN 480–481)

He's talking about the natural version of the motive here, but one might worry that the same concerns beset redirected public interest.

The worry for my account has two parts. I'll take them in order of how easy I think they are to answer. First, Hume is concerned that the motive of public interest is "too sublime" to be the missing motive. In calling a motive "sublime,"

I take Hume to mean that it is other-directed in a way that may seem accessible to saints or angels, but not to (at least typical) human beings. In defense of my view, I would say that a redirected concern for the public interest is not too sublime because it makes the motive conditional on others adhering to the rules as well, and, while I abide by the rules (in the name of public interest), I also demand that others abide by them, including by respecting my property, keeping promises made to me, etc. A concern for the public interest does not preclude a concern for one's own share in it. After all, we are each members of the public. This motive, conditional as it is, does not strike me as too sublime to be the missing motive. It is not the kind of universal benevolence that Hume dismisses.

The second part of the worry is that people "look not so far as the public interest, when they pay their creditors, perform their promises, and abstain from theft, and robbery, and injustice of every kind." This is the sense in which public benevolence is "too remote" to be the motive that we're after. Here, Hume seems to be raising the concern that public interest can't be the motive because it simply isn't what people (including, we must assume, people with the virtue of justice) think about when they are moved to do as justice demands. And if that's the case, it can't be the motive that renders justice a virtue.

In response, I agree with Hume that when I pay a debt, for example, the explicit thought that drives me is not that this particular act of paying my debt is in the public interest. For one thing, it might not be. That's why public benevolence is a bad candidate. But for another, the explicit thought that drives me is more likely to be simply that the debt is owed (or that the rules require that I pay the debt). That said, I think this is consistent with my interpretation.

The motive is best characterized, at least frequently, as an implicit motive. The explicit thought that produces motivation to pay my debt is that it is owed. But what gives that thought motivational force is my conditional concern for the public interest, shaped by the conventions. For this concern to serve as a motive, all that needs to be true is the following. First, the thought that the debt is owed must be sensitive to whether the rules of the conventions are genuinely in the public interest. Second, enough others must also be adhering to the rules. That the rules are in the public interest need not enter my explicit thoughts. I simply need the thought that the debt is owed to produce motivation that is sensitive to whether the conventions are in fact in the public interest. If I were to discover that the rules were not in the public interest or that other people were not adhering to the rules, then the thought that the debt is owed would not move me (and, in fact, it wouldn't strike me as true).

I should clarify this point about implicit motives with another example. Suppose that I pick my daughter up from school. If asked for my motive, I would probably say that I don't want her to be unsafe or feel abandoned. But that thought was likely not in my head at the moment that I left to pick her up. More

likely, the thought in my head was something like “school’s out.” But this doesn’t mean that my desire for her to be safe and not feel abandoned wasn’t my motive. The question is what gives the thought “school’s out” motivational force? The most plausible answer is my desire for my daughter to be safe and not feel abandoned. If my daughter had stayed home sick, then the thought “school’s out” wouldn’t produce motivation at all because it isn’t, in this instance, connected in the right way to my concern for my daughter’s safety and well-being. Explicit thoughts such as “school’s out” produce motivation only when they are connected in the right way to implicit motives such as my concern for my daughter’s safety and well-being. Similarly, explicit thoughts such as “the debt is owed” produce motivation only when they are connected in the right way to implicit motives such as my redirected concern for the public interest.

With this account of the motive, we can make sense of Hume’s claim that people don’t look as far as the public interest when they pay their debts while at the same time maintaining that redirected public interest is what makes our explicit thoughts motivationally efficacious. This motive is, then, neither too sublime nor psychologically unrealistic to be the motive constitutive of the virtue of justice.

However, another issue is that one of the reasons that Hume thinks that the natural version of the motive of public interest can’t be the motive of justice is that this kind of universal benevolence is too *weak* a motive to counteract the motive of self-interest. According to Hume, “That is a motive too remote and too sublime to affect the generality of mankind, *and operate with any force in actions so contrary to private interest as are frequently those of justice and common honesty*” (T 3.2.1.11 / SBN 480–481, my emphasis). Given that acting justly often requires that we act against our own immediate self-interest, a natural concern for the public interest is not a plausible candidate for the motive of justice. The concern for my account is that even a redirected concern for the public interest, given that it has a similar object as the natural version of the motive, would also be too weak to counteract the motive of self-interest.

Hume is right that most people don’t have the kind of universal benevolence at issue in this passage. And to the extent that people have some concern for everyone else (which Hume thinks they do because of the operation of sympathy [T 3.2.1.12 / SBN 481–482]), this concern doesn’t usually motivate grand acts of self-sacrifice for the betterment of the whole. But the redirected concern for the public interest is, I’ve suggested, different from the universal benevolence that Hume dismisses. We approve of those conventions that overall benefit both the public as a whole and each individual member (including, of course, ourselves). And it is these benefits that motivate the just person’s conformity with the rules of the conventions. Taking a longer-term perspective on these benefits and, of course, approving of one’s character in light of reflection upon one’s

proper concern for these benefits, the just person recognizes that doing as justice demands is not a long-term sacrifice. So the just person does not experience the redirected concern for the public interest as in tension with self-interest, even as self-interest is not what motivates this person's characteristic behavior. For this reason, I don't think that the concern that leads Hume to dismiss the natural version of the motive of public interest applies to the redirected version of the motive.

So far I've been arguing that my proposal is *consistent* with the text, despite some potentially worrisome passages. But is there any *positive* evidence in Hume's writings for my view? There is. My account harmonizes nicely with what is, in this literature, an apparently oft-overlooked passage from Hume's *A Dissertation on the Passions*. In defense of his motivational psychology, Hume argues, as he does in the *Treatise*, that what people are talking about when they talk in terms of acting from reason is action motivated by a calm passion. As an example, Hume writes, "A man adheres to justice from reason; that is, *from a calm regard to public good, or to a character with himself and others*" (DP 5.2: 24, my emphasis). As we learn in the *Treatise*, the calm passions operate in a way that is not always fully accessible to us on introspection: "Now 'tis certain, there are certain calm desires and tendencies, which, tho' they be real passions, produce little emotion in the mind, and are more known by their effects than by the immediate feeling or sensation" (T 2.3.3.8 / SBN 417). There is a scholarly dispute about whether the calm passions could be *entirely* unfelt while still being *passions*, which Hume thinks enter our minds "with most force and violence" (T 1.1.1.1 / SBN 1–2). But the interpretive question that we must ask is whether Hume is talking only about the *feeling* of the calm passions or also their content. Could the content of the calm passions be similarly known more by its effects?

I think that Hume should allow for this possibility (whether he intended to or not). After all, one of his examples of a calm passion is "the love of life" (T 2.3.3.8 / SBN 417), which seems clearly almost always motivationally active but very rarely part of the content of our motivating thoughts. So I think it's reasonable to read Hume as claiming that a redirected concern for the public interest is the motive of the just person (the one that originally renders justice a virtue) despite rarely forming a part of the explicit content of the just person's motivational thoughts. And notice that accepting *this* claim doesn't require me to think that the calm passions are entirely unfelt. I can remain neutral on this issue. A calm regard to the public interest, acting implicitly through thoughts such as "the debt is owed," might very well be *felt* (faintly) despite the just person not explicitly thinking about the public interest when she acts and despite its content not being explicit in the just person's mind at the moment of action.

What should we say about Hume's inclusion of "a regard to...a character with himself and others" as another possible motivation for the just person? I

take it that Hume is referring here to a concern for one's reputation. The most plausible rendering of this motivation as the motive of a just person would be the motivation to cultivate a deserved reputation for being just. But now, of course, the circle argument threatens. This can't be the motive that originally renders justice a virtue because it presupposes that justice is a virtue.¹⁶ A (properly qualified) concern for the public interest, then, seems to be, even according to Hume himself, at least one of the motives characteristic of the just person, and it seems a good candidate, as I've argued, for being one that could originally render justice a virtue.

Let's turn now to the question of whether the motive of a redirected concern for the public interest can be squared with Hume's motivational psychology. Darwall, as I've mentioned, thinks that Hume's account of the motive of justice is inconsistent with his motivational psychology. Of course, Darwall has a particular account of the motive of justice (the motivational state of rule-regulation) that does not look to a further end than obeying the rules. This explains Darwall's skepticism because Hume's motivational psychology commits him to the view that motivation depends on prospective pleasure or the avoidance of prospective pain and regulating one's conduct by rules for its own sake does not look to any prospective pleasure or the avoidance of any prospective pain. This problem isn't unique to Darwall's account of the motive as we saw with Sayre-McCord, but its force depends on an account of the motive that either looks to no further end than obeying the rules or looks to an end that has no clear connection with prospective pleasure or pain.

My account makes the motive one that looks to prospective pleasure, in particular, the prospective pleasure of each person affected by the operation of a convention, so it seems at first glance, at least, to render Hume's account of the motive of justice consistent with his motivational psychology, despite Darwall's concern.

Darwall might push back in two ways. First, while my view takes the just person's actions to be motivated by prospective pleasure, the pleasure in question is mostly the pleasure of other people. Darwall, however, thinks that we should read Hume's motivational psychology *egoistically* (1995: 294). On this view, the only pleasure that matters for my motivation to act is *my own* pleasure. If the egoistic reading of Hume's motivational psychology were correct, then a redirected concern for the public interest might be thought to have very little

16. Jason Baldwin proposes a developmental view according to which there are different approved motives for acts of justice. In the final stage of development, the motive is "love of reputation" (2004: 294). Lorraine Besser-Jones has a similar view according to which a "redirection of pride, prompted and enabled by the conventions of justice, [which] tracks the development of a concern for reputation" is the motive (2006: 272). For the reasons just given, I don't think that this can be the motive that originally renders justice a virtue.

motivational force because my own part in the public interest is quite small and everyone else's interest wouldn't move me.

There are two reasons not to attribute this egoistic reading of his motivational psychology to Hume. First, none of the passages that Darwall cites in favor of his reading (T 1.3.10.2, T 2.3.3.3, T 2.3.9.7 / SBN 118–119, 414, 439) restricts Hume's claim that pleasure or the avoidance of pain is necessary for motivation to an individual's own self-directed pleasure or pain. Second, Hume explicitly rejects egoism in the second *Enquiry* (EPM 5.17 / SBN 219).

The second way in which Darwall might push back on my claim that my account of the motive is consistent with Hume's motivational psychology is to insist that the pleasure or pain necessary for motivation must be realized in each particular action in order to motivate that action. If this were right, then my account of the motive, which takes the prospective pleasure to be the pleasure realized through the successful operation of conventions and not pleasure realized through each particular action in accordance with the demands of justice, would still be inconsistent with Hume's motivational psychology.

Lorraine Besser-Jones disputes this way of understanding Hume's motivational psychology. On her view, as long as "the will underlying this pattern of actions [i.e., abiding by the demands of justice] aims at a concrete end-state of desire satisfaction" (2006: 262), then it doesn't matter if the desire in question (on my view, the desire to promote the well-being of each person affected by the convention) is not satisfied by each particular action. I agree with Besser-Jones, and I would add that there's some textual evidence that this is Hume's view. He writes, "Men often counter-act a violent passion in prosecution of their interests and designs: 'Tis not therefore the present uneasiness alone, which determines them" (T 2.3.3.10 / SBN 418). His claim here is that we are capable of overriding an immediate desire in the course of pursuing a longer-term project. In the case of justice, his view, I believe, is that we can override the immediate desire for material gain (or any other immediate concern) in the pursuit of the project of securing the well-being of each person through the successful operation of a convention.¹⁷

This concludes my response to Darwall. But at this point, my view faces an important question about Hume's discussion of the sensible knave, a character that Hume introduces in the *Enquiry Concerning the Principles of Morals* who questions why it is in his interest strictly to abide by the demands of justice when there are occasions on which he could secure some benefit for himself without undermining the value of the conventions. First, a little setup. Hume distinguishes between two motivational questions about justice, the question of what

17. I take these bits of evidence to be burden-shifting, rather than definitive, responses to Darwall.

motivates us originally to form and abide by the conventions and the question of what motive constitutes the *virtue* of justice. Hume is perfectly clear about the former motive (T 3.2.2.24 / SBN 498–500), which is a redirection of self-interest. The latter motive is the subject of this paper. Some commentators, perhaps most notably David Gauthier (1979) and Annette Baier (1991; 2010), see continuity between these motives. On their view, redirected self-interest is not only the motive of the establishment of the conventions, it also transforms into a virtue once we notice that the conventions, sustained by self-interest, deliver benefits to the public. However, as Gauthier recognized in a later paper (1992), this interpretation of the motive constitutive of the virtue of justice runs into a problem with the sensible knave. Why would *self-interest* recommend uniform adherence to the conventions when there are clearly at least imaginable cases in which one could advance one's own interests by committing an injustice without thereby undermining the stability and functioning of the conventions? Wouldn't the better policy to adopt be one of typical adherence to the conventions while occasionally violating their rules when one could do so without significant risk of detection?

Enter Don Garrett. Don Garrett, who shares Darwall's view that the motive of justice is an unyielding (though, I've allowed, conditional) desire to regulate one's conduct by the rules of the conventions, thinks that his view provides Hume with resources to respond to the sensible knave. On Garrett's view, the motive of justice is the product of the adoption of a policy of uniformly abiding by the rules. In the genealogical story that Hume tells about the development of the conventions, we must imagine the people who are developing a motivational orientation towards the conventions as, at this stage, quite socially inept. They have only recently started interacting with others. They have done so enough to recognize that living without conventions of, e.g., property, is highly inconvenient and destructive of the benefits of social interaction, but they are hardly schemers at this point. Rather, they are desperate to resolve the inconveniences of life without the conventions, so they would adopt a policy of uniform adherence. The option of adopting a policy of general adherence while taking advantage of the exceptions, Garrett thinks, wouldn't be likely to occur to such socially inexperienced people (2007: 269).

But once they have adopted this policy, they notice that uniform adherence to the rules delivers benefits to the public, which generates moral approbation for the policy of uniform adherence and serves as a way of sustaining the policy of uniform adherence even in the face of temptation to break the rules, given that moral approbation of the policy both involves *disapprobation* of breaking the policy and is connected with a concern to maintain one's reputation as a morally good person.¹⁸ And moreover, those who commit injustice and are caught

18. For clarity, Garrett thinks of the motive of justice as a policy that we adopt in light of our recognition that the conventions of justice deliver benefits to the public. In this way, one might

are subject to social condemnation and punishment, both unpleasant experiences. Thus, for Garrett, Hume can respond to the sensible knave by claiming that adoption of the policy of uniform adherence, constituted by an unyielding commitment to abiding by the demands of the conventions, is the better policy to adopt (2007: 270–271).

One point of clarification before turning to the relevance of this discussion for my view. Though Garrett's interpretation gives Hume something to say to the sensible knave, the success or failure of Garrett's view (or any other view) doesn't hinge on whether the knave can be brought to accept that adopting the policy of uniform adherence is genuinely in his interest. After all, the knave could say in response to Garrett, "Sure, I can see why if I disapproved of violating the rules of justice, I would have some further interest-based reason to adopt the policy of uniform adherence, but I don't, so I still don't see why uniform adherence would be in my interest. And moreover, I'm a *sensible* knave, so I'm unlikely to be caught." Hume himself recognizes that some people's moral psychologies are sufficiently warped that they simply can't see why acting morally would be in their interest (EPM 9.23 / SBN 283). That alone wouldn't show that justice isn't a virtue or that such a person wouldn't have a moral obligation to abide by the demands of justice.¹⁹

What is the problem for my view supposed to be? The problem is that Garrett is able to give a response to the sensible knave because he counsels the knave not to look beyond the rules of the conventions for guidance about whether to abide by them. The knave keeps wondering if abiding by the rules on this particular occasion is in his interest. But Garrett can say, "No, you're making a mistake from the get-go. What's really in your interest is adopting the policy of uniform adherence and then looking no further than the rules of the conventions for action guidance." The issue is that on my view, the motive of justice, like the knave's motive of interest, looks beyond the rules of the conventions and, in the case of my view, queries their effects on the public good.

think that my view is quite similar to Garrett's, given that we both think that the just person is motivated to abide by the demands of the conventions and that the benefits of the conventions play a role in explaining why. This apparent similarity masks a deep disagreement. For Garrett, once we adopt the policy of uniformly abiding by the rules of the conventions, the benefits that the conventions deliver to the public drop out of the agent's motivation altogether. As a result, though the benefits that the conventions deliver play a role in the causal history of the generation of the motive of justice, on Garrett's view, they play no role in the motive of justice itself. This is a significant difference between his view and mine.

19. I'll note here that those who think that enlightened self-interest is the motive of justice could dispute this point. And it is, in fact, the case that Gauthier gives his error-theoretic interpretation of the Humean virtue of justice precisely because self-interest can't motivate the range of just actions, as the sensible knave, he thinks, demonstrates. But I don't think that the self-interest account is the right account.

That said, and as I've already suggested in my discussion of Darwall, the fact that the motive constitutive of the virtue of justice, on my account, looks out to the effects of the conventions on the public good does not show that what just people are motivated *to do* is to ignore the conventions and do whatever they think is in the public interest on each occasion. The just person is motivated uniformly to abide by the conventions because the just person recognizes that the conventions are the way that we have settled on how to promote the public good. And so given that my view, no less than Garrett's, counsels uniform adherence to the conventions, the challenge of the sensible knave, at least on this point, does nothing to favor one view over the other.

But I believe that Hume's response to the knave actually makes my view more attractive than Garrett's. Hume recognizes that the knave may not be able to be brought to see that abiding by the demands of justice is in his interest. The reason, according to Hume, is that "Inward peace of mind, consciousness of integrity, a satisfactory review of our own conduct; these are circumstances very requisite to happiness, and will be cherished and cultivated by every honest man, who feels the importance of them" (EPM 9.23 / SBN 283). The knave doesn't feel the importance of them. But those who do feel the importance of them feel that they are important, presumably, precisely because they recognize the value that their morally good conduct has.

For Garrett, however, the motive of justice is alienated from the value that justice has, on Hume's view. Hume is clear that what drives our approval of justice is that it is in the public interest: "*a sympathy with public interest is the source of the moral approbation, which attends that virtue [viz., the virtue of justice]*" (T 3.2.2.24 / SBN 498–500). My account makes the motive harmonize with what drives our approval of the virtue. We approve of the virtue because it leads people to abide by the rules of the conventions, which are themselves in the public interest. And the motive constitutive of the virtue, on my account, is a (suitably qualified and conditional) concern for the public interest. But a concern for public interest falls out of the picture in Garrett's account of the motive. As a result, it isn't clear why Garrett's just person *should* have a "satisfactory review of [her] own conduct." Her concern to abide by the demands of justice, after all, is not a concern for what makes justice valuable. This, I believe, is a reason to favor my view over Garrett's.

4. Conclusion

Contemporary moral and political philosophers most commonly think of justice as a property of rules, laws, or social institutions. They less frequently think of

justice as a virtue of character.²⁰ I think it's a shame that the philosophical tradition has lost this focus, in part because it played such a prominent role in the political thought of philosophers such as Aristotle, Aquinas, and Hume. But in addition, this focus has led theorists of virtue and the virtues to lose out on vigorous philosophical theorizing about one of the central virtues.

Now, I freely admit that the account of the virtue of justice I have developed in this paper depends on a lot of complex and interrelated arguments in Hume's moral philosophy. I don't pretend to have defended the account I have given as the right account of the virtue of justice. I do, however, think that 1) the account I have offered is a serious contender for being the one to solve Hume's problem of the missing motive, and 2) the background features of Hume's moral philosophy on which my account relies are not easily dismissed. The account I attribute to Hume, then, merits consideration as the correct account of the virtue of justice.

Acknowledgements

I would like to thank Geoff Sayre-McCord and several anonymous referees for helpful feedback on earlier drafts of this paper.

References

- Abramson, Kate (2015). What's So 'Natural' about Hume's Natural Virtues? In Donald C. Ainslie and Annemarie Butler (Eds.), *The Cambridge Companion to Hume's Treatise* (333–368). Cambridge University Press.
- Baier, Annette (2010). *The Cautious Jealous Virtue: Hume on Justice*. Harvard University Press.
- Baier, Annette (1991). *A Progress of Sentiments: Reflections on Hume's Treatise*. Harvard University Press.
- Baldwin, Jason (2004). Hume's Knave and the Interests of Justice. *Journal of the History of Philosophy*, 42(3), 277–296.
- Baron, Marcia (1982). Hume's Noble Lie: An Account of His Artificial Virtues. *Canadian Journal of Philosophy*, 12(3), 539–555.
- Barry, Brian (1989). *Theories of Justice: A Treatise on Social Justice, Vol. 1*. University of California Press.
- Besser-Jones, Lorraine (2006). The Role of Justice in Hume's Theory of Psychological Development. *Hume Studies*, 32(2), 253–276.
- Bloomfield, Paul (2011). Justice as a Self-Regarding Virtue. *Philosophy and Phenomenological Research*, 82(1), 46–64.
- Cohon, Rachel (2006). Hume's Artificial and Natural Virtues. In Saul Traiger (Ed.), *The Blackwell Guide to Hume's Treatise* (256–275). Blackwell Publishing.

20. Some exceptions include Bloomfield (2011), LeBar (2014), and Schmidtz & Thrasher (2014).

- Cohon, Rachel (2008). *Hume's Morality: Feeling and Fabrication*. Oxford University Press.
- Darwall, Stephen (1995). *The British Moralists and the Internal 'Ought' 1640–1740*. Cambridge University Press.
- Darwall, Stephen (1993). Motive and Obligation in Hume's Ethics. *Noûs*, 27(4), 415–448.
- Fieser, James (1997). Hume's Motivational Distinction between Natural and Artificial Virtues. *British Journal of the History of Philosophy*, 5(2), 373–388.
- Garrett, Don (2007). The First Motive to Justice: Hume's Circle Argument Squared. *Hume Studies*, 33(2), 257–288.
- Gauthier, David (1992). Artificial Virtues and the Sensible Knave. *Hume Studies*, 18(2), 401–428.
- Gauthier, David (1979). David Hume, Contractarian. *The Philosophical Review*, 88(1), 3–38.
- Harris, James A (2010). Hume on the Moral Obligation to Justice. *Hume Studies*, 36(1), 25–50.
- Hume, David (1757/2007). *A Dissertation on the Passions and A Natural History of Religion*. Ed. Tom L. Beauchamp. Oxford University Press.
- Hume, David (1751/1975). *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Ed. L.A. Selby-Bigge, revised by P.H. Nidditch. Oxford University Press.
- Hume, David (1739–1740/1978). *A Treatise of Human Nature*. Ed. L.A. Selby-Bigge, revised by P.H. Nidditch. Oxford University Press.
- Kelahan, Emily (2018). Naturalness and Artificiality in Humean Virtue Theory. *Hume Studies*, 44(2), 249–276.
- LeBar, Mark (2014). The Virtue of Justice Revisited. In Stan van Hooft (Ed.), *The Handbook of Virtue Ethics* (265–275). Routledge.
- Mackie, J. L. (1980). *Hume's Moral Theory*. Routledge.
- Postema, Gerald (1988). Hume's Reply to the Sensible Knave. *History of Philosophy Quarterly*, 5(1), 23–40.
- Rawls, John (1999). *A Theory of Justice: Revised Edition*. Harvard University Press.
- Salter, John (2012). Hume and Mutual Advantage. *Politics, Philosophy, and Economics*, 11(3), 302–321.
- Sayre-McCord, Geoffrey (1995). Hume and the Bauhaus Theory of Ethics. *Midwest Studies in Philosophy*, 20(1), 280–298.
- Sayre-McCord, Geoffrey (2016). Hume on the Artificial Virtues. In Paul Russell (Ed.), *The Oxford Handbook of Hume* (435–469). Oxford University Press.
- Schmidtz, David, and John Thrasher. The Virtues of Justice. In Kevin Timpe and Craig A. Boyd (Eds.), *Virtues and Their Vices* (59–74). Oxford University Press.
- Woozley, A. D. (1978). Hume on Justice. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 33(1), 81–99.