



**David Hume, Paternalist**

Michael Ridge

*Hume Studies* Volume 36, Number 2 (2010), 149-170.

Your use of the HUME STUDIES archive indicates your acceptance of HUME STUDIES' Terms and Conditions of Use, available at

<http://www.humesociety.org/hs/about/terms.html>.

HUME STUDIES' Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the HUME STUDIES archive only for your personal, non-commercial use.

Each copy of any part of a HUME STUDIES transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

For more information on HUME STUDIES contact

[humestudies-info@humesociety.org](mailto:humestudies-info@humesociety.org)

<http://www.humesociety.org/hs/>

## David Hume, Paternalist

MICHAEL RIDGE

*Abstract:* A standard worry about Hume's account of justice is that it leaves those who are most vulnerable outside the circumstances of justice. An equally standard reply is that those who are so vulnerable as to fall outside the scope of justice need not thereby fall outside the scope of morality altogether, because on Hume's account we will often have duties of humanity to treat vulnerable creatures decently. It is not clear that this reply is adequate, for given the apparent priority of justice over natural virtues like those of humanity, it is not clear that duties of humanity provide enough protection for the weak. This paper identifies another problem with Hume's reply: if those who are extremely vulnerable are nonetheless rational and fully capable of autonomous judgment about how to live, then Hume's theory still delivers the wrong *sorts* of protections for them. In particular, it is very plausible to suppose that it would be immoral to engage in paternalistic interference with the decisions of such weak but rational agents about how to live, as long as they are not thereby harming or wronging anyone else. Treating such rational but vulnerable agents paternalistically seems unjust, but Hume's account cannot vindicate this intuition. Indeed, Hume not only cannot explain why such paternalism is unjust, he seems forced to conclude that we will often be *obligated* by duties of humanity to engage in such paternalism. For Hume seems committed to the unpalatable conclusion that morality speaks unambiguously in favour of paternalistic interference in such cases as long as we can be reasonably sure that the intended beneficiaries really will be made better off. In this paper, I develop and press this new objection from paternalism against Hume's account of the circumstances of justice.

## Introduction

Were there a species of creatures intermingled with men, which, though rational, were possessed of such inferior strength, both of body and mind, that they were incapable of all resistance, and could never, upon the highest provocation, make us feel the effects of their resentment; the necessary consequence, I think, is that we should be bound by the laws of humanity to give gentle usage to these creatures, but should not, properly speaking, lie under any restraint of justice with regard to them, nor could they possess any right or property, exclusive of such arbitrary lords. Our intercourse with them could not be called society, which supposes a degree of equality; but absolute command on the one side, and servile obedience on the other. Whatever we covet, they must instantly resign: Our permission is the only tenure, by which they hold their possessions: Our compassion and kindness the only check, by which they curb our lawless will: And as no inconvenience ever results from the exercise of power, so firmly established in nature, the restraints of justice and property, being totally *useless*, would never have place in so unequal a confederacy. (EPM 3.18; SBN 190–91)<sup>1</sup>

It seems fair to say that this passage does not represent David Hume's finest moment. Indeed, it is hard to think of a single other passage from Hume's moral philosophy which is both still taken seriously at all and, at the same time, subject to as much sustained critical attention as this one.<sup>2</sup> In spite of the impressive body of work criticizing this aspect of Hume's theory, it is subject to an interesting and important objection that has not previously been discussed. The objection is interesting, in part, because it arises out of reflection on the implications of Hume's main reply to the more standard objections. The standard worries arise out of the idea that Hume's theory implausibly excludes the powerless from the charmed circle of justice. The standard Humean reply is that at least they fall within the scope of beneficence and kindness. In this paper, I argue that in one respect this actually *magnifies* the force of the original objection. Justice can serve not only to protect one's welfare; it can and should also serve as a bulwark against the paternalistic interferences of do-gooders. Insofar as the species of creatures Hume imagined are owed duties of beneficence and kindness yet are not afforded the protections of justice, such paternalistic interventions may sometimes be not only morally permissible on Hume's theory but also morally *required*. Thus, the worry about the absence of justice, now understood in its anti-paternalistic role, is magnified rather than reduced by the presence of duties of beneficence. In this

paper, I develop this objection from paternalism and explain how it presents an important challenge for Hume's moral theory.

Before developing the objection, I must first put the offending passage in its proper context. After all, objections arising out of this passage would not be so interesting if the passage itself expressed ideas which Hume could easily jettison without undermining the main themes of his larger project. Moreover, the fact that there is no analogue to this passage in Hume's discussion of the so-called "circumstances of justice" in the *Treatise* might well make a sympathetic reader wonder just how deeply committed Hume was, or anyway should have been, to the ideas expressed in it. I, therefore, begin by putting the passage in the context of Hume's wider theory of justice and, indeed, his moral theory more broadly (section 1).

Having put this passage into its proper context and argued that it would not be easy for Hume to reject its main ideas, I move on to a discussion of some of the more familiar worries about his view (section 2). A standard worry is that Hume's account leaves those who are most vulnerable outside the circumstances of justice. An equally standard reply, given by Hume himself, is that those who are so vulnerable as to fall outside the scope of justice need not thereby fall outside the scope of morality altogether. In particular, on Hume's account we will often have duties of humanity to treat such creatures decently.

Whatever one thinks of this problem taken on its own terms, there is another problem arising out of Hume's reply (section 3). For if those who are extremely vulnerable are, nonetheless, rational and fully capable of autonomous judgment about how to live their own lives, then Hume's theory still delivers the wrong *sorts* of protections for them. In particular, it is plausible to suppose that it would be immoral to engage in paternalistic interference with the decisions of such weak but rational agents about how to live their own lives—so long as they are not thereby harming or otherwise wronging anyone else. Treating such rational but vulnerable agents paternalistically seems unjust, but Hume's account cannot vindicate this intuition. Indeed, Hume not only cannot explain why such paternalism could be unjust; he seems forced to conclude that we will often be morally *obligated* by duties of humanity to engage in such paternalism. For Hume seems committed to the unpalatable conclusion that morality speaks unambiguously in favour of such paternalistic interference in such cases so long as we can be reasonably sure that the intended beneficiaries really will be made better off.

The problem of paternalism is different in character from the more familiar problems of exclusion and inequality. The problem of exclusion arises because Hume's theory of justice perversely excludes the vulnerable, when one might have thought the vulnerable stood most in need of its protection. The problem of inequality is that while the vulnerable do get some moral protection on Hume's theory, albeit not the protections of justice, these protections are pitifully weak. The problem of paternalism is different again. Rather than arguing that the

protection afforded the vulnerable by duties of humanity is too weak, which is a sort of quantitative notion, the problem of paternalism is that duties of humanity do not provide protections *of the right kind*. More precisely, the problem of paternalism brings out the fact that there seems to be no space in Hume's theory for reasons of the right kind for not interfering with the vulnerable but rational. Sometimes we should not interfere with such agents not because non-interference is best for them (it may clearly not be) but because we respect them and their ability to decide for themselves how to live, what risks to take, how to rank different values, and so on. Indeed, the problem is worse than this. For the very premise Hume invokes to defend his view against the problem of exclusion is that we at least owe duties of humanity to such helpless but rational creatures. It is precisely because we owe them duties of humanity but no duties of justice that Hume seems forced to say that odious forms of paternalism are actually sometimes obligatory.

## 1. Hume and the Circumstances of Justice

In this section, I will put the passage with which I opened into its proper context and thereby explain why Hume cannot easily shed the commitments he takes on in that passage. My strategy here is to argue that Hume understands both the origin and merit of justice in terms of its utility *qua* mutual advantage and then to argue that this understanding makes it very difficult for him to shed the problematic commitments. There are, in principle, strategies of extension which would allow Hume (or a Humean, anyway) to avoid these commitments, but I argue (in section 2) that these strategies are also problematic. Before proceeding, I should address the fact that the offending passage has no analogue in the *Treatise*. After all, many scholars consider the *Treatise* to be the more sophisticated and impressive work and for this reason might try to discount this passage as a matter of interpretation. However, as a matter of pure exegesis, this is implausible. For Hume explicitly took the *Enquiries* to reflect his more considered and mature view.<sup>3</sup> Moreover, if I am right that Hume's conception of justice in both the *Treatise* and the *Enquiries* is based on the mutual advantageousness of conventions of justice, then the considerations which drive Hume to take on this commitment in the *Enquiries* are already present in and central to Hume's thinking in the *Treatise*.

The passage with which I opened this paper occurs in Hume's discussion of what later scholars, most notably John Rawls, have called the "circumstances of justice."<sup>4</sup> For Hume, justice itself is, in the first instance, a virtue and, hence, a character trait which garners approval "from a common point of view."<sup>5</sup> Hume understands justice in terms of property and promise-keeping, so justice *qua* character trait and virtue presumably is constituted by a sort of settled disposition to respect people's property rights and keep one's promises for the right

sorts of reasons. Just what count as the right sorts of reasons is part of what is at issue in the discussion of the circumstances of justice. One of Hume's main aims in that discussion is to establish that we approve of justice because of its utility, where its utility is, in turn, essentially mediated by convention. Its utility is essentially mediated by convention not only, or even primarily, in that the utility of any given just action may depend on the existence of background conventions of property. Its utility is essentially mediated by convention in the deeper sense that it is the proper functioning of those conventions itself which is the real source of utility. People's being just and adhering to those conventions is valuable primarily as a means to the maintenance of the relevant conventions, on Hume's account. Hume's view is opposed to those that maintain that the sentiment is "like hunger, thirst, and other appetites, resentment, love of life, attachment to offspring, and other passions, arises from a simple original instinct in the human breast, which nature has implanted for like salutary purposes" (EPM 3.40; SBN 201). Hume takes great pains to enumerate the ways in which the particular details of our property conventions are arbitrary, historically contingent, and heterogeneous, both within any given society at different times and between different societies (EPM 3.39–3.48; SBN 201–04). This, in his view, casts great doubt on the idea that our approbation of justice is due to an "original instinct." This line of argument is also obviously connected to Hume's contention (more emphasized in the *Treatise*) that justice is an artificial virtue rather than a natural one.

At the same time, it is important to Hume that justice is a genuine and important virtue. One might, therefore, worry that the arbitrary and sometimes even whimsical nature of particular property conventions might cast doubt on this. Hume's strategy is to argue that it does not matter much just what property conventions exist as long as they function to coordinate our interactions in the right way. It is in this context that one must see Hume's argument that both the origins and the merit of justice are to be understood in terms of its utility. For this argument will be mediated by the argument for the utility of the relevant conventions, and the argument for their utility must not, in turn, depend on the particular details or content of those conventions. The overall argument is that human beings are much better off having some property conventions rather than none. Indeed, on a standard reading of Hume's view, having some such conventions in place is so important that the marginal expected utility of moving from one set of property conventions to another is almost never great enough to risk undermining the stability of the existing conventions. This last contention is notoriously conservative, and one might reasonably object to Hume's view on these grounds. As objections to Hume on this front have been raised before, though, I shall not here go over that well-trodden ground.<sup>6</sup>

Hume explains both the origin and the merit of justice entirely in terms of its utility and repeatedly asserts that public utility is the “*sole* origin of justice and that reflections on the beneficial consequences of this virtue are the *sole* foundation of its merit” (EPM 3.1; SBN 183). So, there are two distinct ideas captured by the notion of the circumstances of justice, one corresponding to Hume’s descriptive and the other to his justificatory ambitions. The first, descriptive, element of the concept is that the circumstances of justice are those circumstances in which justice conventions are likely to be invented or maintained by normal human beings. Hume moves from this rather vague concept of the circumstances of justice—as those circumstances in which justice is likely to arise and have some merit—to the more specific and controversial conception of those circumstances in which justice is useful, identifying justice’s value with its instrumental value. This conception is still not sufficiently precise, though; it immediately raises the questions “Useful as compared to what?” and “Useful to whom?” How we answer these questions is important; if justice must simply be useful to *someone* as compared to *some* conceivable alternative, then the circumstances of justice will be very broad indeed and much broader than if they required justice to be useful to everyone when compared to *every* conceivable alternative.

Hume is fairly clear that the answer to the first question, “useful as compared to what?” is the baseline of cooperation without conventions of justice in small groups that we find in the Humean state of nature.<sup>7</sup> This comports well with the conservative strand in Hume’s thought discussed above. The point is to argue that the advantages of some stable conventions of justice, when contrasted with our having no such stable conventions, is so great as to justify adhering to them quite strictly and not trying to modify them for any marginal gains in utility due to the danger of destabilizing those conventions. For this reason, Hume criticizes egalitarian movements like the “Levellers” for not recognizing the dangers of trying to reform the status quo (EPM 3.2.4; SBN 193).

Just how Hume answers the second question, “Useful to whom?” is more controversial. In my view, the text and arguments best support the answer, “everyone to be bound by the relevant conventions.” To this extent, I follow David Gauthier’s reading of Hume’s theory of justice as being a form of justice as mutual advantage.<sup>8</sup> Since Gauthier has defended this interpretation at length and, in my view, quite plausibly, I shall not here reinvent the wheel. I shall, however, make some remarks about why his reading seems so plausible. One piece of textual evidence can be found in the *Treatise*, where Hume remarks, “whatever may be the consequence of any single act of justice . . . the whole system of actions, concurr’d in by the whole society, is infinitely advantageous to the whole, *and to every part*” (T 3.2.2.22; SBN 498; my emphasis). The addition of “to every part” can plausibly be read as indicating that each individual bound by the conventions

reaps some benefit from the existence of those conventions and the activities which constitute them.

Hume's discussion of justice in the international sphere also provides evidence for the mutual advantage interpretation. In this sphere also, Hume links moral obligations of justice with utility and goes so far as to hold that in circumstances in which justice is less useful it is also less obligatory. In particular, he suggests that the obligations of justice are less stringent between nation-states than they are between individuals. This apparently is because nation-states are more capable of subsisting on their own than individuals:

But nations can subsist without intercourse. They may even subsist, in some degree, under a general war. The observance of justice, though useful among them, is not guarded by so strong a necessity as among individuals; and the *moral obligation* holds proportion with the *usefulness*. All politicians will allow, and most philosophers, that reasons of state may, in particular emergencies, dispense with the rules of justice. . . . But nothing less than the most extreme necessity, it is confessed, can justify individuals in a breach of promise, or an invasion of the properties of others. (EPM 4.3; SBN 206)

If by "utility" Hume meant something more like "aggregate utility" in the utilitarian's sense, then it is hard to see why the rules of justice would be any less stringent between nation-states than between individuals. Hume argues that norms of justice are less binding in the international sphere precisely because each individual nation-state has less to gain from the existence of such international justice conventions.

Hume further supports his claim that both the origin and the merit of justice can be found in its utility by discussing the various circumstances in which justice would have no utility and arguing that in such circumstances justice would never have arisen in the first place. It follows from Hume's more general theory that outside the circumstances of justice, conventions of property and promising would have no merit and hence, could not provide a basis for any moral obligations. Again, see the discussion of justice in the international sphere. Because conventions of property and promising which were extended to include the helpless but rational creatures would not be mutually advantageous, the default Humean view must be that any such creatures would fall entirely outside the scope of justice. No doubt, this is why Hume's own explicit view was that any such creatures would not be able to reap the benefits associated with such conventions. I shall later (in section 2) explore some ways in which Humeans might try to get beyond this "default position" for the rational but helpless creatures, perhaps by invoking some of Hume's own ideas about the universalizing tendencies of general rules, the

progress of our sentiments, and the role of sympathy in broadening our conventions. First, though, I must say more about Hume's account of the circumstances of justice themselves.

An important motif in Hume's discussion of the circumstances of justice is that extremities of various kinds would render justice useless. Hume argues that under conditions of extreme abundance and of unlimited benevolence, justice would be useless, in the first case because dividing property would serve no purpose when property is unlimited, and in the second case because justice would be superfluous so long as everyone acted from unlimited concern for the flourishing of others. Further, Hume urges that under conditions of extreme scarcity, as well as under conditions of extreme selfishness, justice could serve no purpose, because in such situations justice would inevitably give way to "the stronger motives of necessity and self-preservation" (EPM 3.8; SBN 186).<sup>9</sup>

Most importantly for present purposes, Hume suggests in the *Enquiries* that the circumstances of justice would not hold between us and creatures who are incapable of all resistance and unable to make us feel the effects of their resentment. Keeping in mind Hume's conception of justice, it is not too difficult to see why he is driven to this doctrine. Consider, in turn, the descriptive and the normative elements of the circumstances of justice. Given that justice is that cautious, jealous virtue, which owes its origin to self-interest, it is no surprise that a species of creatures utterly incapable of having an impact on our self-interest will not be included in the scope of justice. History supports Hume's descriptive thesis on this point. Even the relatively powerless have frequently been denied the protection of justice. The absolutely powerless presumably would fare much worse.

With respect to the normative element, it is no surprise that Hume held this view since he thinks justice owes its merit solely to utility, where utility is understood as *mutual advantage*. If these creatures cannot have any influence on the interests of the dominant group, then it might well not be advantageous for the members of the dominant group to extend the protection of justice to such creatures. Hence, including such creatures within the scope of justice would, at least on the face of it, have no merit of the sort relevant to Hume's project. Unless some special story can be told about how those parties who can influence one another's interest would benefit by extending the conventions of justice to some such powerless group (a strategy of extension), justice, as Hume understands it, simply will not apply to them. Unless some such strategy of extension can be made to work, then, the conception of justice as mutual advantage commits Hume to the problematic exclusion of the helpless but rational creatures in the passage on which I shall focus in this paper. I now turn (in section 2) to some of the classic responses to the objections to the offending passage, some of which explore what I am here calling the "strategy of extension." I argue that these strategies are problematic and that it is not so easy for someone with Hume's core commitments to

shed the view he takes of the helpless but rational creatures discussed in the key passage. Having further bolstered the thesis that Hume cannot easily shed this view, I develop the main objection to Hume's view that I want to explore in this paper—the objection from paternalism (in section 3).

## 2. The Problem of Exclusion and the Problem of Inequality

It is not hard to see how Hume's view can seem problematic. On a plausible understanding of justice, one of its primary functions is to protect the weak from being exploited by the strong, whereas on Hume's account, justice extends no protection at all to the utterly weak and powerless; if they are incapable of resisting and making the effects of their resentment felt, then the "necessary consequence," Hume claims, is that we do not "lie under any restraint of justice with regard to them." One would have thought, however, that it is *particularly* unjust when the strong exploit the weak, and the ability of the strong to condition themselves or be conditioned so that nothing the weak can do will make them feel the effects of their resentment does nothing to mitigate this injustice at all. Hume's critics have been quick to raise this objection. Thomas Reid, for example, pressed this point, arguing that justice is especially important, rather than without merit, in conditions of extreme inequality. Reid argues that Hume is mistaken in claiming that justice would be "totally useless" in the envisioned circumstances, arguing that, "surely to be treated with justice would be highly useful to the defenceless species he here supposes to exist."<sup>10</sup> Reid's objection should highlight an important feature of Hume's view; for the circumstances of justice to hold among creatures, it must be useful to *each* of them, at least compared to the baseline of no juridical cooperation. So, Reid's objection rests on an oversimplified reading of Hume's view. Hume's account may be implausible, but it is not vulnerable to this very quick rejoinder.

Gerald Postema has usefully distinguished two related worries one might have about Hume's view.<sup>11</sup> The first is the *problem of exclusion*, the worry that Hume's view altogether excludes certain individuals from the scope of justice who should fall within that scope. The second is the *problem of inequality*, the worry that while Hume's view may include the *very* weak but not utterly powerless in the scope of justice, the protection it affords them will be pitifully weak. Brian Barry, one of the most outspoken modern critics of Hume's doctrine, has shifted the emphasis of his case against Hume away from the problem of exclusion and toward the problem of inequality.<sup>12</sup> However, while the problem of inequality is a deep one for many *Humean* views, the problem of exclusion strikes more at what is essential to Hume's own view. This is because the problem of inequality only applies to views that make the content of justice depend upon relative bargaining power, as contemporary contractarian accounts like David Gauthier's do.

Whether Hume himself is committed to a contractarian model may reasonably be contested; all that one is committed to in embracing Hume's account of the circumstances of justice is that relative bargaining power, *at the limit*, has an impact on whether one is included within the scope of justice *at all*. One could accept this element of Hume's view but then have some non-contractarian model of the principles that should be adopted to govern all those who do fall within justice's scope, and that model might be completely insensitive to differences in bargaining power.<sup>13</sup> In fact, as I noted above (in section 1), Hume's conservatism in this area consists in his emphasis on the value of some conventions of justice over none, and the view seems to be that this provides a strong presumption in favor of deferring to the justice conventions in place in one's society *whatever their content* so long as they are better than the baseline of the Humean state of nature. For this reason, Gauthier himself argues quite convincingly that Hume's theory is a coordination theory rather than a bargaining theory: "Hume conceives of the problem of selecting among rules as one of coordination, rather than bargaining. Bargaining, the typical contractarian device, is a relatively costly procedure for reaching agreement, suitable only when our differential preferences among possible conventions are strong in comparison with our interest in the selection of some convention rather than none."<sup>14</sup> The fact that Hume insists on the arbitrariness of our actual property conventions at some length underscores the plausibility of Gauthier's interpretation on this point.

The problem of exclusion is a deeper one. To be clear, the worry is *not* primarily a worry about which actual groups should have been and should be included in justice's scope. Rather, it is a worry about the grounds upon which any group or individual ought to be included. Hume tells the wrong story about the merit of including others within justice's scope. So even if, for any *actual* group or individual, Hume's view entails they should have been (or be) included in the scope of justice if and only if we intuitively think they should be included in its scope, his theory would still make their inclusion depend upon the wrong sorts of considerations. The problem Hume faces is quite general. In many *imaginable* circumstances, young children, highly vulnerable tribes, the handicapped, the elderly, and distant future generations will all be difficult to include in the circumstances of justice in virtue of their utter powerlessness. Since distant future generations cannot resist us or make the effects of their resentment felt for the trivial reason that a cause must precede its effect, they may present the most difficult problem for Hume's view.<sup>15</sup> The real worry is that Hume makes the answer to the question, "should this group be afforded the protections of justice?" contingent on the wrong kinds of considerations. To paraphrase Jeremy Bentham, the relevant question is not "can they make their resentments felt?" but "are they capable of rationally restraining their behavior in accordance with reasonable conventions of property and promising?" For justice, unlike kindness, does perhaps presuppose some measure

of reciprocity. It is presumably for this reason that Hume is right that we do not owe duties of justice to nonhuman animals, though we do of course owe them duties of kindness and the like.

It is here important to note two features of Hume's hypothetical helpless creatures. First, they are rational, and second, it is presupposed that they can feel resentment. Hume explicitly says that they are rational, and the claim that they "cannot make their resentments felt" is most naturally read as implying that they may have resentments. If Hume's view were, instead, that the creatures he imagined were incapable of feeling resentment in the first place, then the way he put his point would be highly misleading at best. These points together undermine an alternative reading of the passage proposed by Arthur Kuflik.<sup>16</sup> Kuflik argues that the best interpretation of Hume in this passage is one which relies on the idea that justice for Hume involves following rules. He suggests that, like nonhuman animals, the creatures Hume imagined would be incapable of following rules as sophisticated as those which govern our conventions of property and promising. This, in turn, explains why resentment is relevant here. For resentment is not simply anger, but anger informed by one's recognition of an infraction against oneself of a certain sort, for example, by an infraction of widely accepted conventions of property. Insofar as a species of creatures is unable to grasp and follow rules or feel resentment when those rules are violated in ways that infringe their rights under those rules, the practice of extending those rules to that species of creatures is problematic.

The problems with this reading are fairly clear. First, as Kuflik himself notes, this reading does not fit at all well with the emphasis on ideas of relative power in the passage. Kuflik suggests that the only way to capture those aspects of the passage is to attribute to Hume the view that only rough equality of power is sufficient to fall within the circumstances of justice. Kuflik argues that this reading is not only uncharitable but also fits very poorly with other things that Hume says, indeed, things he says shortly after he describes the helpless creatures. Fortunately, there are ways of understanding the relevance of power here that do not go so far as requiring anything like rough equality of power. Instead, we can take the text at face value: what is necessary is a power, all right, but not a power that is sufficient for rough equality. All that is needed is enough power to make one's resentments felt. Small children can certainly make their resentments felt: it is very hard to ignore a child having a screaming tantrum because he sees his treatment as "not fair!" Yet small children and their parents (and other adults) do not stand in anything like "rough equality." Kuflik's argument here relies on a false dichotomy; there are other ways of understanding the role that power is playing in this passage.

Secondly, and more decisively, Kuflik's interpretation rests on the idea that it is the inability of the creatures in question to follow rules and feel resentment

that explains why Hume excludes them from the charmed circle of justice. Yet, as I have already indicated, the passage simply does not bear this reading. The creatures are rational, after all, which would seem to be sufficient for following simple rules of property and promise-keeping. For example, insofar as they are rational, such creatures can choose to keep whatever promises they make. Indeed, it is presumably precisely because actual nonhuman animals are not rational in the relevant sense that Hume felt it necessary to “go hypothetical” here, thus illustrating that on his view, mere possession of rationality alone is not sufficient for someone to fall within the circumstances of justice. If his point were simply that justice requires a certain measure of reason, then he could have made his point perfectly well with ordinary nonhuman animals (whose rationality he here leaves as an open question). Admittedly, Hume does say that the creatures are so weak both in body *and mind* as to be unable to make their resentments felt. In my view, this sort of weakness does not entail an inability to follow simple rules of property and promising, though. Instead, it just means that the creatures are not clever enough to overcome their physical weakness and make us feel their resentments through some subtle emotional manipulation, or the like. The lack of this sort of manipulative and psychological cleverness does not entail an inability to follow simple rules.

Finally, of course, Hume’s way of putting his point strongly suggests that the creatures he imagines do feel resentment even though they cannot make us feel the effects of that resentment. As Kuflik himself argues, however, resentment is itself not simply anger, but anger informed by some conception of a rule of some kind (either of justice or of some other sort of moral rule, perhaps) having been violated in a way that somehow grounds a complaint by oneself in particular. This feature of Hume’s discussion also strongly undermines the proposed reading.

A natural Humean response to the problem of exclusion is to provide some indirect way to include within the scope of justice the helpless but rational creatures that Hume imagines, Hume’s own official views notwithstanding. To be theoretically interesting, though, the strategy must be recognizably Humean. The most interesting move here would be to argue that our sympathy with such creatures and our recognition of the ways in which they could benefit from inclusion within our conventions of property and promising might naturally lead us to extend our conventions so as to include them. Moreover, as Hume frequently points out, the details of our conventions of property and promising are themselves often the result of general rules, which have a universalizing tendency. For example, Hume presumably would include the friendless elderly within the circumstances of justice, but one might debate the extent to which the friendless elderly are capable of making the effects of their resentment felt. Our natural sympathy with such individuals and the tendency to form general rules that go beyond the paradigm

cases motivating the rule in the first place might well explain why our conventions tend to have this scope.<sup>17</sup>

This reply is probably the best one going for the Humean, but it does not entirely meet the objection. First, the reply still makes the moral status of the rational but helpless creatures dependent on the range of our sympathy. It is perhaps not for no reason that Hume's own example is couched in terms of another *species* of creatures.<sup>18</sup> For a start, this is obviously a harder case than one involving helpless members of our own species. In that case, our sympathy is very easily evoked; sympathy, like charity, begins at home, after all. Moreover, the more different the species is from our own the less clear it is that our sympathy will naturally extend to the members of that species. If the species of creatures were vaguely humanoid or perhaps just cute (think of how people respond to squirrels as opposed to rats in virtue of their bushy tails), then our sympathy might well easily be evoked. This is not the hard case, though, and perhaps Hume implicitly had in mind a species of creatures much more deeply different from us and one that we might find aesthetically ugly or even disgusting. Think of a species of intelligent but helpless slug-like, rat-like, or serpent-like creatures, for example. No matter how angelic their behavior to one another may be and how clear it may be to us that they are every bit as rational as ourselves, our sympathy may simply not stretch that far. Nonetheless, it is intuitive that such creatures could make reasonable claims of justice upon us. The Humean might simply bite the bullet here, but this is not without cost.

Note, by way of contrast, that a Kantian moral theory, with its emphasis on the *a priori* sentiment of respect for the moral law, which impartially applies to all rational agents as such, can do much better on this front. Whether a broadly sentimentalist theory could do better than Hume's is a harder question.<sup>19</sup> Perhaps a view like Simon Blackburn's contemporary "quasi-realism" could be developed in a way that is at once broadly Humean and sentimentalist, while preserving the idea that the protections of justice depend neither on the ability to make one's resentments felt nor on the ability to evoke our sympathy.<sup>20</sup> After all, Blackburn's quasi-realism is meant to remain neutral on first-order questions, and the extent to which the protections of justice are dependent on such considerations is, within his framework, a first-order issue. Moreover, quasi-realism is, if successful, meant to provide a Humean vehicle for preserving our pre-theoretical ideas of the mind-independence of our moral obligations. Insofar as duties of justice's independence from the scope of our sympathy is a form of mind-independence, this seems like a promising way to preserve much of what is insightful about sentimentalism without encountering the problems arising out of Hume's own view.

Secondly, Hume might also reasonably worry about how realistic conventions of justice that are based entirely on the sympathy of fiduciaries could be.

By hypothesis, the helpless creatures will not, by themselves, be able to enforce their rights: they are powerless to make the effects of their resentment felt. It will only be in virtue of the assistance of those of us who are inclined to promote their interests that the rights afforded to them by courtesy of extending our conventions of property and promising could be enforced. This would mean that in their case justice would have to depend on such benevolent impulses being suitably widespread and strong or on a few enforcers having sufficient power. Hume, of course, is on record arguing that benevolence, even in our own case, is too weak a motive to oppose self-interest, and for this reason he argues that benevolent motives cannot provide the origin of justice. If that point is sound at all, then it must be even sounder when we turn to a deeply alien species of creatures with whom our sympathy may be more severely limited in its scope and depth. After all, defending the property rights of such helpless creatures will not be without substantial costs, as those who try to exploit them will often be powerful. Given how important it was to Hume to provide a theory of justice that is suitably realistic, this seems to be a serious worry about the appeal to sympathetic fiduciaries as a solution to the problem of exclusion.

In this section, I have explored the traditional debate over the problem of exclusion and concluded that the problem is a serious one for Hume and for Humeans who endorse Hume's account of the circumstances of justice. My point has not been to settle this debate but to put it on the table in order to raise a new but related objection. I now want to open up a new front against Hume's account, albeit one which builds naturally on the problem of exclusion. Indeed, it is really a new reason to worry that Hume's appeal to duties of humanity should not blunt whatever force the problem of exclusion has in the first place, for when unchecked by duties of justice, duties of humanity can sometimes require us to help other competent and fully rational agents against their better judgment about how to lead their own lives. Such duties can in some contexts smack of a morally dubious paternalism, or so I shall argue.

### 3. Hume's Paternalism

Consider a paradigmatic case of justified paternalism, a parent's forcing a young child to do something the child does not want to do because it is really, uncontroversially in the child's best interest, all things considered. Intuitively, the reason parents may legitimately intervene in these ways is because young children are not yet competent to judge what is in their best interests. We tend to think, by contrast, that we cannot force competent adults to do what we see to be in their best interests *simply* because we see that it is in their best interests. In other words, whether paternalism is justified with regard to someone should be a function of that person's rational competence. Unfortunately, by making a

person's inclusion in the scope of justice depend upon her ability to resist and make the effects of her resentment felt, Hume cannot accommodate this widely held moral intuition. This is because constraints on paternalistic intervention would have to be constraints of justice rather than constraints of humanity. Call this the "problem of paternalism."

An example should help make the point. Suppose we are on a spaceship traveling through the galaxy, and our spaceship has transporters like those featured on *Star Trek*. In fact, a scenario similar to the one I describe actually arises in *Star Trek: The Next Generation*, in the episode entitled, "The Ensigns of Command" (season 3, episode 2).<sup>21</sup> We come upon a small population of creatures who are quite rational. We might even suppose that in many ways they are more intelligent than we are, in spite of our much more advanced technology. Perhaps they have devoted their superior intellects to philosophy and literature rather than to technological advances. In any event, we have just discovered that their planet is in the midst of a severe ecological crisis, one that will make the lives of the inhabitants quite miserable. We explain the situation to the inhabitants and also make it clear that we could transport them to another planet that is *very* much like their own but that does not face any such ecological crisis. Moreover, we explain this to them in a way that makes their situation quite clear. Nonetheless, they all voluntarily decide to remain on their home planet, in spite of this information. For them, their emotional attachment to their home planet, with all its idiosyncrasies and its distinctive history, is too great for them to favor abandoning it, no matter what the costs. We are quite certain, though, and with good evidence, that they will in the long run be better off on this other planet. Their nostalgia for their home planet will fade with time, and any such psychological costs will be vastly outweighed by the advantages of avoiding a severe ecological crisis and all its accompanying ills (starvation, disease, and so on). We therefore "beam" them, against their wills, to this other planet and provide them with the means to start a stable civilization there.<sup>22</sup>

Let us stipulate that because we have superior technology, the members of this alien species are completely unable to resist us or make the effects of their resentment felt.<sup>23</sup> Also, let us stipulate that they have not fortuitously fallen within our pre-existing justice conventions (as they might have if they were humanoid and our conventions were put in terms of "all rational humanoids," for example). On Hume's view, it seems there simply is no issue of justice between us and this "species of creatures." Hume can only appeal to the natural virtues to criticize what we have done. Such an appeal will be unsuccessful.

Our actions are perfectly compatible with the demands associated with the natural virtues. Were we, for example, *inhumane* in transporting these creatures to another planet? Quite the opposite; we spared them enormous physical and psychological trauma and caused them only some short-lived anxiety and, in our

view, overly sentimental sadness. We took great pains to do for the members of this species what we genuinely and reasonably believed was in their best interests, which is hardly a case of behaving inhumanely. Forcing them to move was no more inhumane than forcing one's four-year old child to move to a new town because you are unemployed and must move to find work and put food on the table. Indeed, a very good case can be made out that *not* forcing them to move would be inhumane. Moreover, note that if we drugged the creatures in question in a way that would alleviate their anxieties and resign them to their new lives away from their old planet, or, perhaps even altered their memories, this would hardly help to address our anti-paternalistic worries. Such interference with the lives, and even the identity, of these creatures would make our actions more, rather than less, objectionable. Yet considerations of benevolence in and of themselves do not obviously speak against such measures.

The basic point is simple. The most powerful objection to what we have done to this alien species is *not* that it was inhumane but that it was unjust. We treated these beings just like children, incapable of deciding what was best for themselves, when they were in fact quite competent to do so. We did not respect their right to continue living on *their* planet if they made a competent and informed decision to do so.<sup>24</sup> This is, however, just the objection that Hume could not raise to what we have done. Hume, therefore, seems committed to seeing what we have done as an instance of justified, indeed perhaps morally obligatory (as a matter of kindness, say), paternalism. This can be made even clearer by reflecting on the fact that if these creatures had fallen within the circumstances of justice but had been considerably *less* rational and not as competent to determine their own best interests, Hume would, or at any rate could, have said it was unjust to force them to relocate. Hume makes the justification of paternalism especially sensitive to differences in power. Intuitively, though, the justification of paternalism should instead be sensitive to differences in intellectual and rational capacities.

Hume could of course just respond, "So much the worse for your anti-paternalistic intuitions," in much the same way that some contemporary utilitarians try to "out-smart" their opponents by accepting the supposed *reductio* of their view as a surprising truth rather than an absurd implication of a false view.<sup>25</sup> Hume himself was, after all, no contemporary liberal, and our strictures against paternalism may not have moved him. Fair enough, but I do not claim to have made a case that given his own historically embedded assumptions, Hume must revise his view. That, I suspect, would be rather anachronistic. My point is more for those who take Hume's moral theory in general and his account of the circumstances of justice in particular to be a plausible view worthy of consideration in contemporary debates about morality and justice. In spite of the historical gap between us and him, much of what Hume has to say does, indeed, still seem relevant and important to

contemporary moral and political philosophy. My aim has been to hear Hume as speaking to us with some potential relevance and see whether or not his theory is defensible even from a perspective that gives more weight to strictures against paternalism than Hume himself might have.

Another and more interesting line of reply maintains that my objection depends on an uncharitable gloss of Hume's conception of our duties of humanity and beneficence.<sup>26</sup> In my *Star Trek* example, those on the *Enterprise* who beam the inhabitants of the planet away without their consent might be seen by Hume as arrogant do-gooders, who may not even really or, in any event, deeply care about the people they are "saving." Moreover, their judgments about what would be best for the inhabitants of that planet may be colored by their own preferences. Hume himself seems keenly aware of the dangers of projecting one's own preferences and values onto others. Recall, for example, that in the essay, "The Sceptic," Hume repeatedly inveighs against the tendency (especially amongst philosophers) to project one's own values on to others and to take an overly narrow view of the forms a good life might take. For example, Hume remarks that, "it cannot reasonably be doubted, but a little miss, dressed in a new gown for a dancing-school ball, receives as compleat enjoyment as the greatest orator, who triumphs in the splendor of his eloquence, while he governs the passions and resolutions of a numerous assembly."<sup>27</sup>

The objection is a fair one so far as it goes, but because it can be met by slightly modifying the example, it does not pose a deep and serious challenge to my argument. The worry that those on the *Enterprise* do not really care about the people they forcibly transport can easily be dealt with. Since the story is an imagined one, we can simply stipulate the facts. So, let us stipulate that the concern of the crew is genuine and heartfelt. In my view, this should not lead us to abandon the verdict that their paternalistic interference is morally objectionable, though it may make their characters more sympathetic.

The more interesting worry is that their concern, while genuine, is misguided. This is where Hume's observation in "The Sceptic" is relevant. The worry is that a deeper appreciation of the perspective of the inhabitants of the planet would reveal that it is better for them to remain where they are, the enormous costs of doing so notwithstanding. There are three points to make about this interesting defence of Hume's view. First, while Hume does sound a reasonable and important scruple, it does not follow that a suitably impartial and careful investigator cannot accurately and reliably assess what course of events will make someone more or less happy. We can again stipulate that the crew in our example are careful in just these ways, and, indeed, that they are right that the inhabitants *will* in the long run be happier if transported. Indeed, part of the point of the robust liberal anti-paternalistic intuitions mobilized by the argument is that such interference with the choices of competent and rational adults is problematic even when we are

right about what would be best for them. A complication here is that there could be children involved, for whom paternalism would be appropriate. However, we can stipulate that all the children of this species were recently killed by a plague which killed only the young.

Secondly, it is also an important methodological maxim in gauging what will make people happy or unhappy to bear in mind the malleability of their preferences. At least, this is true with human beings, and I stipulate it is true (and known to be true) of the hypothetical alien species in my example. For example, many carnivores who accept the moral arguments for vegetarianism but refuse to change their dietary habits are firmly of the view that they would be miserable without meat in their diet. In fact, though there may be rare exceptions, this belief generally grossly underestimates the malleability of people's gustatory preferences. Firmly held values are perhaps harder to change than gustatory preferences, and I may have overestimated the severity of the limitations on such changes, but the literature on adaptive preferences makes an impressive case for the view that people tend to be very adept at such modifications.

Thirdly, there are certain disvalues which are not variable between people. In particular, physical pain, starvation, exposure to the elements, the constant struggle against an unforgiving and relentlessly harsh natural environment, loss of loved ones, and the like are pretty universally experienced as very negative. Granted, even here there are exceptions, for example those who relish the challenge of extreme adversity, but these exceptions are very rare indeed (and again, I stipulate that my imagined creatures are like human beings in this respect as well). The point is that in the forced relocation example, the costs of remaining on their home world are very steep in these ways. So, even if the crew of the *Enterprise* has to some extent not properly evaluated the positive preferences of the helpless aliens, it could be that even on a more fair evaluation the costs of their staying are so high that they would be better off moving to the other world to which they are in fact forcibly relocated.

Indeed, unless the defender of Hume denies that people's reflective and informed judgments about what would be best for them can ever be misguided, and clearly misguided to a suitably situated external observer, an example with the right form can be constructed. Notwithstanding Hume's reasonable scruples about not projecting one's own values onto others, such cases are clearly possible. Insofar as anyone with robustly anti-paternalistic moral intuitions will still rebel at the idea of forcible relocation in such a case, the objection stands.

## Conclusion

I have argued that Hume's account of the circumstances of justice is vulnerable to the "problem of paternalism." This problem goes beyond the more familiar

problems of exclusion and inequality by emphasizing the special *character* of justice and contrasting it with the character of benevolence when the latter is unchecked by any sense of justice. Hume himself might not have been much moved by the liberal sentiments behind this objection and might have just followed the argument where it led. However, my point has not been to refute Hume without questioning norms which might have been more prominent in his own day. I have instead tried to see how far Hume's account might still be relevant to moral and political philosophy today. Even within a contemporary framework, a hardy Humean could insist that we give up our anti-paternalistic intuitions on the strength of a suitably Humean theory. It is hard to know how to assess such a move without a wholesale evaluation of Hume's theory, but, in my view, we should resist it. In any event, such a move would constitute a cost of the view, one that needs to be taken seriously.

Finally, though, I do not think these objections force those of us with liberal intuitions about paternalism to abandon the sentimentalist approach to moral philosophy that Hume espouses. For Hume's account of the circumstances of justice is not essential to his larger sentimentalist project or to sentimentalism in general. The problem emerges out of Hume's commitment to understanding justice in terms of utility, where utility in turn is understood in terms of mutual advantage. This approach to justice, though, is in no way mandatory for the sentimentalist. Indeed, in my view, Adam Smith's moral theory is precisely a sentimentalist one, which avoids what is implausible about Hume's own theory, at least when it comes to the circumstances of justice. That, however, is a story for another day.

## NOTES

Many thanks to Geoff Sayre-McCord, Michael Smith, the editors of *Hume Studies* (Corliss Swain and Saul Traiger), and two anonymous referees for helpful comments and suggestions.

1 References to the *Enquiry* are to David Hume, *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Clarendon Press, 1998), abbreviated "EPM" in the text and cited by section and paragraph numbers, followed by page numbers in *Enquiries concerning Human Understanding and concerning the Principles of Morals*, ed. L. A. Selby-Bigge, 3rd ed., revised by P. H. Nidditch (Oxford: Clarendon Press, 1975), abbreviated "SBN" in the text.

2 I am here putting to one side some of the more quaint and sexist passages in Hume's corpus.

3 A point made by David Gauthier. As he reminds us, Hume went so far as to call the *Treatise* "the juvenile work," going on to indicate that the latter work "may alone

be regarded as containing his philosophical sentiments and principles.” See David Gauthier, “David Hume, Contractarian,” *The Philosophical Review* 87 (1979): 3–38.

4 Cf. John Rawls, *A Theory of Justice*, 1971 (Cambridge: Harvard University Press).

5 For a useful discussion of the role of the general (or “common” as it is sometimes put) point of view in Hume’s moral philosophy, see Geoffrey Sayre-McCord, “On Why Hume’s ‘General Point of View’ isn’t Ideal—and Shouldn’t Be,” *Social Philosophy and Policy* 11 (1994): 202–28.

6 For some useful discussion, see D. Livingston, “Hume’s Conservatism,” *Hume Studies* 21 (1995): 151–64.

7 Here I follow Baier’s analysis in “The Laws of Nature.” See Annette Baier, *A Progress of Sentiments* (Cambridge, MA: Harvard University Press, 1991), 220–54.

8 See Gauthier, “David Hume, Contractarian.”

9 See Brian Barry, “Circumstances of Justice and Future Generations,” in *Obligations to Future Generations*, ed. Brian Barry and R. I. Sikora (Philadelphia: Temple University Press, 1978). See also C. Hubin, “The Scope of Justice,” *Philosophy and Public Affairs* 9 (1979): 3–24.

10 See T. Reid, “Essays on the Active Powers of the Mind,” in T. Reid, *Philosophical Works*, ed. Sir William Hamilton, 8th ed., vol. 2 (1895; repr., Hildesheim, 1967), 2: 660.

11 See Gerald Postema, “Making Resentment Felt: Hume on the Environment of Justice,” unpublished manuscript.

12 See Barry, “Circumstances of Justice,” and Brian Barry, *Justice as Impartiality* (Oxford: Clarendon Press, 1995). Barry later came to the view that the problem of exclusion, as opposed to the problem of inequality, is not a real one for Hume.

13 Postema also makes this point. See Postema, “Making Resentment Felt,” 37.

14 See Gauthier, “David Hume, Contractarian,” 23–24.

15 This is especially worth noticing in the context of Hume’s theory of justice, which gives such pride of place to property as the paradigm institution of justice, since current institutions of private property may have a profound impact on the well-being of distant future generations.

16 In Arthur Kuflik, “Hume on Justice to Animals, Indians and Women,” *Hume Studies* 24 (1998): 53–70. Thanks to Saul Traiger and Corliss Swain for calling Kuflik’s article to my attention.

17 Thanks to an anonymous referee and the editors of *Hume Studies* for very useful dialogue on this point.

18 Here I part company with Joyce Jenkins and Robert Shaver, who argue in their “Mr. Hobbes Could Have Said No More,” in *Feminist Interpretations of David Hume*, ed. Anne Jaap Jacobson (University Park, PA: Penn State University Press, 2000), 137–55, that Hume is committed to the view that women, who of course are members of our own species, fall outside the circumstances of justice. They argue, quite correctly, that this thesis would commit Hume to a morally outrageous sort of paternalism with

respect to women, substituting “gallantry” for justice. Although I admit that Hume was a product of his times and sexist in a number of ways, I am not convinced that he is committed to the thesis that women fall outside the circumstances of justice. A full defence of this position would go beyond the scope of this essay. However, I will make three quick points. First, endorsing such a strong thesis would make Hume not only a product of his time but also reactionary even by the standards of his time. For even in Hume’s day, it would have been thought unjust to break a promise to a woman, and unmarried women were not entirely without property rights. Second, the textual support for this reading of Hume is, even by the lights of Jenkins and Shaver, not unambiguous. Third, one would have thought that if Hume held such a view, he would not have felt the need to advert to a merely hypothetical species of creatures in the passage. Indeed, as Jenkins and Shaver are aware, Hume goes on just after the helpless but rational creatures passage to indicate that women are typically able to break the oppression of their “lordly masters” through “insinuation, address and charms.” A fairer criticism of Hume would be that his conservatism about the content of the rules of justice (noted above in the text) would make him much less motivated to argue for a more fair or equitable arrangement of property rights for women, who still had to give up all of their property upon marriage.

19 Thanks to an anonymous referee for pressing me to discuss this issue.

20 See, for example, Simon Blackburn, *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993).

21 “The Ensigns of Command,” *Star Trek: The Next Generation*, season 3, episode 2. Writer: Melinda Snod; Director: Cliff Bole. First aired 30 September, 1989.

22 It is worth noting that although I used a rather fantastical example to make it clear that we have creatures that are sufficiently helpless relative to us, the general issue I’m raising is far from strictly academic. The People’s Republic of China and other despotic regimes have frequently engaged in massive forced relocations. Though their motives were far from humane, as I am imagining us to be in this thought experiment, they do serve as a vivid reminder that, from a practical point of view, it is important to ask whether forced relocation of the utterly helpless is just. A better example, suggested to me by an anonymous referee, is the forcible relocation of people in Newfoundland and Labrador from remote locations to more central ones in order to provide them government services like health care and education. Again, even though the motives here are laudable in some respects, anyone who has liberal sentiments against paternalism will find this sort of forcible relocation deeply problematic.

23 One might wonder why we even gave them the illusion of making up their own minds about whether to stay if we really were so dismissive about the grounds they would offer for doing so. We might have done so, however, not because we thought their reasons were good but because we thought they would be happier on their new planet if they had autonomously chosen to abandon the old one.

24 Note that an intuitive way of understanding the case is one in which the collective *property* rights of the inhabitants of this planet to *their* planet are being violated. Given the Humean conception of justice as limited to property and promising, this is worth emphasizing.

25 The reference of “out-Smart” is to J. J. C. Smart, who is fond of such bullet-biting moves in the context of his defense of utilitarianism.

26 Thanks to Corliss Swain and Saul Traiger for pressing me on this point.

27 David Hume, *Essays: Moral, Political and Literary* (Indianapolis: Liberty fund, 1985), 166.