



Utilitarianism with a Humean Face

Elizabeth Ashford

Hume Studies Volume 31, Number 1, (2005) 63 - 92.

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

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

Utilitarianism with a Humean Face

ELIZABETH ASHFORD

Introduction

There is a long-standing debate over whether or not Hume's moral theory¹ should be viewed as some version of utilitarianism.² Among opponents of a utilitarian reading, many contrast the subtlety and psychological plausibility of Hume's account of morality with what they take to be utilitarianism's failure both to capture the complexity of morality and to be suited to the nature of human beings.³

Geoffrey Sayre-McCord is, in my view, one of the most incisive and forceful advocates of this position.⁴ He interprets Hume as an exponent of what he calls the "Bauhaus theory" of ethics. The Bauhaus theory avoids a commitment to a single overarching moral principle by which to solve all moral questions. On this view, the virtues are each well-suited to solving a particular problem which we find collectively salient when we take up the general point of view, "and different problems may demand dramatically different, sometimes incompatible solutions" (282). Sayre-McCord contrasts this with what he considers the overly systematic nature of utilitarianism, which identifies traits as virtues solely on the grounds of their tendency to realise or promote well-being. In addition, he contrasts the Bauhaus theory, which "portrays morality's demands as grounded in the expansive interests and affections, generous and not, of real people looking for . . . solutions to problems they face," with what he takes to be utilitarianism's psychological implausibility.

The question of whether Hume's moral theory should be viewed as a version of utilitarianism will, of course, largely depend on how utilitarianism is defined. It is important to separate utilitarianism's key tenets from claims that have come to

Elizabeth Ashford is Lecturer in Moral Philosophy, University of St. Andrews, Edgecliffe, The Scores, St Andrews, KY16 9AL, United Kingdom.
e-mail: ea10@st-and.ac.uk

be associated with the theory but that do not follow from these tenets. My discussion of whether Hume's moral theory is utilitarian will focus on Sayre-McCord's argument that "no version of utilitarianism . . . sits comfortably as an interpretation of Hume." One of the strengths of Sayre-McCord's argument is that he bases it on the most uncontentious and ecumenical account of utilitarianism's core tenets. He argues that there are three ways in which a theory may fail to qualify as utilitarian, and that Hume's theory fails on all three counts:

- (i) a theory might rely on neither actual nor expected utility as the determinants of evaluation; or (ii) it might not take into account all the consequences for everyone affected; or (iii) it might not take utility to be a single measure of interpersonal value.

I will argue that Hume's account of morality adheres to the first two tenets that Sayre-McCord rightly identifies as definitive of utilitarianism, welfarism and impartiality. This means that Hume's theory is a version of what Thomas Scanlon calls "philosophical utilitarianism," according to which "all that counts morally is the well-being of individuals, no one of whom is singled out as counting for more than the others."⁵ By showing that Hume's account is a version of philosophical utilitarianism I hope to demonstrate that a moral theory's adherence to these two fundamental tenets of utilitarianism is compatible with its having considerable complexity and psychological plausibility.

I will then argue that Hume also takes welfare to be "a single measure of interpersonal value," and so also adheres to what Sayre-McCord takes to be essential to the third tenet of utilitarianism. I will argue, *contra* Sayre-McCord, that Hume's account of the moral point of view is based on the interests of all, rather than the interests of each *seriatim* (as on contractualist accounts) or the individual self-interest of the participants (as on Hobbesian contractarian accounts).

Admittedly, though, Hume shows no commitment to maximizing the net sum total of welfare. Such a commitment has come to be viewed as definitive of classical utilitarianism. In fact, as I will argue, Hume does not seriously address the problem of interpersonal conflicts of value, and assumes that the interests of each and of all will generally coincide. However, I will suggest that other early utilitarians too did not adequately face up to the problem of interpersonal trade-offs. I will also argue that while maximizing the net sum total of welfare is an obvious step from the first two tenets of utilitarianism,⁶ it is in fact in deep tension with the most plausible interpretation of these tenets, and moreover the interpretation that Hume, Bentham, and Mill, along with many subsequent utilitarians, had in mind. Furthermore, I will argue that one of the few passages in Hume that does address interpersonal trade-offs closely resembles some key passages in Bentham, that are explicitly concerned with the distribution of welfare. These passages in

Bentham are in tension with those in which he advocates simply maximizing the net sum total of welfare.

I will suggest, then, that Bentham's commitment to maximizing the net sum total of welfare may have resulted from his failure to adequately address interpersonal trade-offs and to consider the other options within impartial welfarism for determining what constitutes the general good. As I will argue, the passages in Bentham that express a concern for the distribution of welfare suggest an alternative and more appealing account of the general good, which follows more naturally from the first two tenets of utilitarianism. Moreover, I will suggest that Hume's model of the impartial sympathetic spectator, or at least a natural development of it, implies this alternative account of the general good.

Utilitarianism has a rich traditional of philosophical thought, yet is often viewed in isolation from that tradition. As a result, the view of what utilitarianism is often bears little resemblance to the way in which it was conceived by its original proponents. However, several of the claims that arise in modern formulations of utilitarianism are not entailed by the key tenets of the theory. Moreover, in my view many of the problems associated with utilitarianism result from these claims, and do not result from the core tenets. This suggests that examining the early versions of utilitarianism may be a particularly fruitful and suggestive way of examining what the key tenets of the theory do and do not entail and how they are best interpreted. As I will argue, Hume's moral theory is especially instructive in this respect. In the case of first two tenets, I will argue that Hume's account adheres to them while avoiding many of the problems faced by modern formulations. I will then argue that Hume does not adhere to the third tenet of utilitarianism as it has come to be understood, but that his account of the moral point of view suggests a conception of the general good that points to an alternative and more credible direction in which utilitarianism might have developed. Since, as I will argue, this conception of the general good follows much more naturally from a plausible understanding of the first two tenets of utilitarianism, and since it is expressed in some key passages in Bentham, it may in fact have a stronger claim to be labelled as "classical utilitarian" than does a commitment to maximizing the conglomerate net sum total of welfare.

I will now turn to Sayre-McCord's arguments against reading Hume as a utilitarian. His arguments are based on a discussion of where Hume's theory stands with respect to each of the three tenets he identifies as definitive of utilitarianism, and I will discuss these arguments in turn.

1. Welfarism

According to welfarism, the single ultimate point of morality is to promote well-being and to reduce suffering. Hume's moral theory gives pride of place to the virtues. Several passages explicitly claim that the ultimate point of the virtues is

to promote well-being. For example, towards the end of the *Enquiry* he describes Virtue as declaring “that her *sole* purpose is, to make her votaries and all mankind, during every instant of their existence, if possible, cheerful and happy” (EPM 9.15; SBN 279, my italics). He continues: “The sole trouble which she demands, is that of just calculation, and a steady preference of the greater happiness.”

In line with this, his classification of the virtues as those traits that are useful or agreeable to their possessor or to others, is fully welfarist. Traits that are useful to others count as virtues because of their tendency to promote others’ well-being. Traits useful to their possessor (such as prudence) are morally valuable because the agent’s happiness is important in itself, as part of the general good. In so far as the agent’s pursuit of her self-interest *conflicts* with the interests of others, Hume holds that it is far from estimable, and describes it in the language of mere egoism.

Traits that are agreeable to their possessor include cheerfulness and greatness of mind, and traits that are agreeable to others include wit and good manners. The distinction between these traits and traits that are useful is grounded in the fact that the former are immediately enjoyable, “valued for the immediate pleasure, which they communicate” (EPM 7.29; SBN 260), either “to the person possessed of them” or to others. They are therefore distinguished from useful traits by being constitutive as opposed to instrumental means to well-being.

Against a welfarist reading, Sayre-McCord argues that according to Hume “we rely not on actual or expected utility in evaluating characters (whether a specific character of someone or a character type) but primarily on an appeal to their usual consequences under standard conditions” (288), and that he therefore fails to qualify as utilitarian. However, it is important to draw a distinction between Hume’s account of the evaluation of character types, which he takes to be stable motivational traits, and his account of the evaluation of the behaviour of particular virtuous agents on particular occasions. Hume takes the status of character types as virtues to be grounded solely on the actual promotion of utility: the virtues are identified as those stable motivational traits that are most useful for human beings in general to have. His analysis of character types is therefore fully welfarist. Our appraisal of particular agents, on the other hand, is directed at whether or not they possess such traits, and is independent of the actual consequences of those agents’ behaviour; clearly, an agent could possess those motivational traits that are the most useful ones for human beings in general to have but be precluded by his particular circumstances from actually promoting welfare. On Hume’s view, then, while the moral appraisal of individual agents is directed at whether they instantiate virtuous character types and is not dependent on the actual consequences of their behaviour, the identification of virtuous character types themselves is grounded on their actual utility.

Moreover, Hume offers important utilitarian reason for taking the focus of moral appraisal to be the consequences of character types, rather than the

consequences of particular agents' behaviour. Hume argues that traits come to be established and passed down as virtues through society's accumulated experience of the kind of traits that generally tend to promote well-being, given the circumstances that obtain in society. The conception of the virtues is therefore informed by a broad experience of human nature, and embodies the collective wisdom of generations. Society's characterization of various virtues and vices gradually changes, when we collectively learn to reassess traits' impact on well-being. For example, "Giving alms to common beggars is naturally praised; because it seems to carry relief to the distressed and indigent: But when we observe the encouragement thence arising to idleness and debauchery, we regard that species of charity rather as a weakness than a virtue" (EPM 2.18; SBN 180). (The empirical assumptions underlying Hume's argument here are highly questionable, especially in the absence of adequate work opportunities and an adequate safety net for those unable to work; but his argument is based entirely on what he takes to be our collective experience of traits' actual impact on welfare.)

The moral appraisal of the virtues, then, arises over time and within a social setting. Because this appraisal is a gradual and collective process it is directed at the consequences of character types, rather than the consequences of particular virtuous agents' behaviour on particular occasions. Likewise, acting virtuously tends to involve the agent's being motivated to adhere to inherited social rules that have been passed down through society in virtue of their general tendency to promote well-being, rather than setting out *de novo* to apply the standard of utility to their particular circumstances.

The importance of this collective wisdom is illustrated with Hume's discussion of the virtue of justice, which describes the dangers of individual agents' attempts to apply standard of utility directly to a particular situation, and the importance of inherited general rules based on society's collective wisdom. Acting justly involves complying with "general, inflexible" established property laws. Given that the distribution of property is unequal, compliance with the established property laws might seem to thwart needlessly persons' interests, since "wherever we depart from this equality, we rob the poor of more satisfaction than we add to the rich" (EPM 3.25; SBN 194). Hume argues, though, that individuals' schemes to change the rules of justice (as was attempted by, for example, the "levellers"), have in fact proved "impracticable," and "extremely pernicious to human society." Since people are unequal in talent and diligence, an equal distribution would quickly lead to inequality anyway. The authority needed to enforce strict equality would therefore have to be so invasive and coercive as to be tantamount to tyranny. Moreover, without a differential incentive to greater productivity, too little would be produced, and so the poor would themselves be among the losers in a system of enforced equality. Over time, society learns the destructive consequences of such unrealistic idealism.

This example illustrates the way in which individuals' attempts to set out *de novo* to apply the standard of utility to raw data are likely to be based on an insufficient knowledge of factors such as human nature and are therefore prone to dangers such as utopianism, with "pernicious" results. As Hume puts it, "a rule, which, in speculation, may seem the most advantageous to society, may yet be found, in practice, totally pernicious and destructive" (EPM 3.23; SBN 193). Society's conception of the virtue of justice is based on a collective experience of human nature, built up over many generations. The virtue of justice involves conforming to inherited social rules that embody this collective wisdom. (Again, while the empirical assumptions underlying Hume's endorsement of rigid compliance with the existing property laws are questionable, his endorsement is grounded solely on the claim that the existing laws do in reality promote welfare better than alternative schemes, given the collective experience they embody.)

The virtue of fidelity illustrates another utilitarian reason for taking the appraisal of agents to be directed at whether they possess those traits that are most useful for human beings in general to have, rather than focusing on the consequences of particular agents' behaviour on particular occasions. The rules of fidelity have arisen entirely because of their usefulness for human beings. The status of fidelity as a virtue is solely dependent on its ensuring "the combination of parents for the subsistence of their young. . . . Without such a *utility*, it will readily be owned, that such a virtue would never have been thought of" (EPM 4.5; SBN 206–7). The virtue of fidelity involves a stable motivation to conform to inflexible social rules against infidelity even when conforming to these rules might not serve this purpose of contributing to child-rearing given the agent's particular situation: although "the laws of chastity" thus "all have a reference to generation," they also apply to "women past child-bearing" (EPM 4.6–7; SBN 207). In this case, the danger with agents' exercising individual discretion over whether to conform to the rules in the light of their particular situation lies with individuals' tendency to favour their own interests in their decisions. It is important for the rules of conduct associated with the virtue of fidelity to be inflexible and apply even to women past child-bearing, because otherwise "women, continually foreseeing that a certain time would bring them the liberty of indulgence, would naturally advance that period" (EPM 4.7; SBN 208). Hume concludes that "*General rules* are often extended beyond the principle whence they first arise": although their origin lies in their role in promoting utility, they apply even in cases in which they no longer do so, and this is important because more flexible laws would in fact end up being abused.

Hume's account in some respects closely resembles contemporary rule and motive utilitarianism, given that it takes the primary focus of moral appraisal to be the consequences of general rules and character types, rather than the consequences of specific agents' behaviour on particular occasions. However, Hume's

account also differs from these contemporary versions of utilitarianism in interesting and important ways. In particular, unlike rule utilitarianism, Hume does not appeal to the consequences of a hypothetical situation of widespread compliance with a set of rules. Hume is concerned solely with the actual consequences of a set of rules. He takes the merit of justice, for example, to be founded on its actually being in the public interest. His account therefore avoids worries about the how ideal world scenarios apply to the actual world. Accordingly, Hume also holds that when the agent's situation diverges so significantly from the standard circumstances to which such rules are adapted that it is sufficiently clear that failure to comply with the rules really will have better consequences, then complying with the rules has no point. For example, in a situation of extreme scarcity or utter villainy, when the standard conditions of justice are no longer in place, there is no point in acting justly.

As I will argue in the following section, Hume's keen awareness of our epistemic and moral limitations also lies behind another respect in which his account differs from rule utilitarianism along with other contemporary versions of utilitarianism: his focus on the improvement of existing conceptions of the virtues and social institutions, rather than on trying to arrive at optimific conceptions of them.

Hume offers a further reason for why our moral approbation of particular virtuous agents is not dependent on the actual or expected consequences of that agent's actions, which marks a significant departure from many subsequent versions of utilitarianism, but is nevertheless compatible with his taking the point of morality to be solely grounded in the promotion of well-being. Moreover, it is in my view a further respect in which his account is more plausible:

If we find, upon enquiry, that the virtuous motive was still powerful over his breast, tho' check'd in its operation by some circumstances unknown to us, we retract our blame, and have the same esteem for him, as if he had actually perform'd the action, which we require of him.

It appears, therefore, that all virtuous actions derive their merit only from virtuous motives, and are consider'd merely as signs of those motives. (T 3.2.1.3-4; SBN 477-8)

Although the identification of traits as virtues is based solely on their tendency to promote welfare, our approbation of virtuous agents *themselves* is unaffected by the fact that circumstances prevent them from performing an action that in fact promotes welfare. This is because Hume takes a person's virtuous character to be a property of *that person*, which consists in his settled motivations. It is independent of the actual consequences of the person's actions on the ground that these consequences are entirely external to the person. Hume holds that an

action assessed solely by its consequences and considered independently of the agent's motivation is just a mechanical claim of cause and effect, which he calls an 'external performance' (T 3.2.1.2; SBN 477). A person's settled motivational traits are what connect him with his actions. Nevertheless, as we have seen, what makes motivations qualify as virtuous is entirely grounded in the actual promotion of welfare. On Hume's view, then, whether particular agents are virtuous is determined by their stable motives, which is a purely internal issue. But what it is for such motives to be virtuous is determined by the consequences of their being generally possessed by human beings.

Virtuous traits therefore elicit two distinct kinds of moral approbation. The first kind is aroused by sympathy with those benefited by the virtues. This leads to the welfarist justification of the virtues, as traits that tend to promote well-being. When we praise a particular moral agent as virtuous, however, our focus is on the agent's character and moral worth, and this is independent of the consequences of that agent's behaviour. Both these kinds of evaluation are described in the following passage:

Where a person is possess'd of a character, that in its natural tendency is beneficial to society, we esteem him virtuous, and are delighted with the view of his character, even tho' particular accidents prevent its operation. (T. 3.3.1.19; SBN 584)

2. Impartiality

Sayre-McCord's second main argument as to why Hume's theory is sharply distinct from any version of utilitarianism is that "when utility matters, it is not utility for *all* who might be affected" (289). In his paper "On why Hume's "General Point of View" isn't ideal—and shouldn't be," he argues that Hume, in his view rightly, holds that "the scope of morality will . . . remain bounded by the actual reach of our sympathetic responses." Against this I will argue that according to Hume, everyone's interests matter from the moral point of view, and matter equally.

The moral point of view is based on the sentiments of humanity and fellow feeling. Hume argues that the scope of these moral sentiments, unlike that of "the selfish passions," is not limited. While the selfish passions "contemplate the greater part of mankind with the utmost indifference and unconcern," "the sentiments . . . which arise from humanity . . . comprehend all human creatures" (EPM 9.7: SBN 273). What qualifies someone as a sensitive moral judge is his having "a warm concern for the interests of our species" (EPM 5.39: SBN 225). Thus no-one's interests are excluded from the scope of the moral sentiments that form the basis of the moral point of view. Hume's moral point of view is, therefore, universalist.

This brings us to the question whether it is also impartial. Hume is certainly very much aware of the influence of our partiality, along with factors such as contiguity, on our felt sympathetic responses. However, he argues that the calm passions which ground the moral point of view correct for the influence of partiality and contiguity on our felt passions. He therefore denies that our moral judgments are bounded by the range of our *felt* sympathetic responses. Our moral judgments are made “without regard to self, or the persons with whom we are more intimately connected” (EPM 5.42; SBN 229), and are oblivious to national boundaries:

Thus a small benefit done to ourselves, or our near friends, excites more lively sentiments of approbation than a great benefit done to a distant commonwealth: But still we know here, as in all the senses, to correct these inequalities by reflection, and retain a general standard of vice and virtue, founded chiefly on general usefulness. (EPM 5.42, footnote 25; SBN 229)

While the violent passions “always plead in favour of whatever is near and contiguous” (T 3.2.7.2; SBN 535), the calm passions are directed at an object’s “real and intrinsic value” (T 3.2.7.2; SBN 534), and comprise “a correction of appearances” (EPM 5.41; SBN 228). When our sentiments are influenced by our partiality we know they do not reflect the “real and intrinsic value” of the benefit or harm, but reflect our personal bias. We therefore know that the fact that we feel less sympathy for distant suffering does not reflect its real importance. For example at T 3.3.1.18 (SBN 583) Hume says:

’Tis seldom men heartily love what lies at a distance from them, and what no way redounds to their particular benefit. . . . Here we are contented with saying, that reason requires such an impartial conduct, but that ’tis seldom we can bring ourselves to it, and that our passions do not readily follow the determination of our judgment.

He goes on to specify that by “reason” here he means “a general calm determination of the passions, founded on some distant view or reflection” (T 3.3.1.18; SBN 583), and unaffected by contiguity and partiality.⁷

The problem of how to reconcile this account of moral judgments as arising from the point of view of an *idealised* observer—to which our actual moral sentiments do not always conform—with Hume’s moral internalism has been forcefully addressed by Elizabeth Radcliffe.⁸ She argues that when we form moral judgments we approximate the moral feelings such an observer would feel; we feel the moral sentiments of an ideal observer to a certain degree, and we extrapolate from these sentiments to our moral judgments. As she points out, this means that

taking up the moral point of view is a matter of degree. Thus, while we can only ever approximate the moral point of view, the moral point of view itself is that of an idealized observer.

As evidence against an impartial reading of Hume's moral point of view, Sayre-McCord appeals to Hume's claim that when it comes to evaluating individuals' characters we focus on a limited circle of those who have commerce with them. However, Hume's explanation of this is that it is psychologically impossible for us to overcome our partiality:

we perceive, that the generosity of men is very limited, and that it seldom extends beyond their friends and family, or, at most, beyond their native country. Being thus acquainted with the nature of man, we expect not any *impossibilities* from him; but confine our view to that narrow circle, in which any person moves, in order to form a judgment of his moral character. (T 3.3.3.2; SBN 602; my italics)

Likewise, he holds that in our judgment of agents "We make allowance for a certain degree of selfishness in men; because we know it to be inseparable from human nature. . . . By this reflexion we correct those sentiments of blame" (T 3.3.1.17; SBN 583).

It should be emphasized that Hume is not giving any intrinsic moral weight to the limited scope of our concern. The language with which he is describing it, with terms such as "limited generosity" and "selfishness," is that of brute egoism. He therefore does not view it as a feature of human nature that the moral point of view itself should accommodate. Rather, the moral point of view from which we appraise agents is fully impartial, and from this perspective, given that a degree of selfishness is an inseparable part of human nature, it limits our moral expectations of agents.

I suggest, then, that while Hume is very much aware of the "selfishness and limited generosity" of human nature, his account of morality itself is fully impartial. As I argued in the last section, it is also welfarist. Nevertheless, his awareness of our moral limitations, and of the limits of our knowledge of how best to promote welfare, gives rise to another important respect in which his account differs from contemporary versions of utilitarianism: he does not appeal to optimific character traits or social institutions, in a technical sense. This would require an impossibly high standard both of knowledge and of selflessness. Rather, he takes the chief role of critical reflection about how to apply the standard of utility to be to decide how existing conceptions of the virtues and social institutions might be improved. This process accommodates both our epistemic and moral limitations.

First, as I discussed in the last section, Hume is skeptical about the feasibility of individuals' attempts to apply the standard of utility directly to their particular

situation. He is particularly skeptical about the possibility of calculating what will maximize social utility in some precise, technical way: "To reduce life to exact rule and method, is commonly a painful, oft a fruitless occupation."⁹ Rather than appealing to what is optimistic, Hume describes how we collectively gradually improve our grasp of what traits and social institutions best promote welfare. He holds that social rules and institutions tend to evolve in such a way as to promote welfare but that they may do so imperfectly, so that there is room for critical reflection about how they might be improved. This critical reflection involves building on the existing experience embodied in them of what tends to promote welfare, rather than setting out *de novo* to apply the standard of utility to raw data. It also tends to involve considering the salient consequences, rather than making precise calculations. For example, the "monkish virtues" of celibacy and self-denial clearly fail to promote welfare, and their status as virtues should therefore be revised. But while Hume takes talk of maximization in some precise or pseudo-precise way to be pointless, the appraisal of traits as virtues is grounded solely on their tendency to promote well-being, and the more welfare they promote the better: "qualities are approv'd of, in proportion to the advantage, which results from them" (T. 3.3.4.11; SBN 612.)

Second, although Hume's account of morality is impartial, he is also, as we have seen, very much aware of the "selfishness and limited generosity" of human nature. Unlike contemporary act utilitarianism, on his view we do not start out with the criterion of the impartial maximisation of well-being as the standard for morally right actions, and label any behaviour that falls short of that as morally unacceptable. Instead he holds that our moral appraisal of agents is based on what we can expect of them, given that "a certain degree of selfishness" is "inseparable from human nature." Improving the existing conception of the virtues and social institutions involves our gradually overcoming some of our moral limitations and extending the scope of our generosity.

Both features are illustrated in Hume's account of the virtue of justice. Property laws evolve because of the crucial role they play in protecting the public good. Hume holds that what is of clear importance for the public good is that there be some "steady and constant" property regulations, so that disputes are kept to a minimum and each person has security. The determination of particular rules involves an inevitable degree of vagueness and is "pretty indifferent":

That there be a separation or distinction of possessions, and this separation be steady and constant; this is absolutely required by the interests of society, and hence the origin of justice and property. What possessions are assigned to particular persons; this is, generally speaking, pretty indifferent; and is often determined by very frivolous views and considerations. (EPM Appendix 3.10, footnote 65: SBN 309)

But while it is fruitless to try to arrive at the optimific set of rules of justice in a precise sense, the overall goal of the rules of justice is to promote well-being as much as possible: the rules are adopted as “best serve the . . . end of public utility” (EPM App. 3.6; SBN 305).

The laws of justice embody a broad experience of the limited degree of generosity that can be feasibly expected, given our current state of moral development. Improving our conception of the virtue of justice is a gradual process, involving a “progress of sentiments” towards expanding the boundaries of our concern: “the boundaries of justice still grow larger, in proportion to the largeness of men’s views. . . . History, experience, reason sufficiently instruct us in the natural progress of human sentiments” (EPM 3.21; SBN 192). (This passage is reminiscent of Peter Singer’s discussion of the expanding circle.¹⁰)

It is worth noting that other key figures in the utilitarian tradition have been equally skeptical about the possibility of making precise welfare calculations. In particular, Mill follows Hume in emphasizing the importance of an awareness of the historical traditions and social settings within which moral sentiments and social rules have evolved, given the breadth of experience of human nature embodied in this process, and he describes attempts to ignore or jettison such traditions and start from scratch to make technical utility calculations as naïve and psychologically shallow.

Moreover, Mill too does not take any action other than the one that is optimific to be morally wrong. He connects the notion of duty with that of the blameworthy, which is confined to certain kinds of actions rather than being applied to all non-optimific actions. For both Hume and Mill, what counts as blameworthy depends on our expectations of others, and these expectations in turn depend on the state of moral development of human nature and society. As Mill argues:¹¹

Inasmuch as every one who avails himself of the advantages of society, leads others to expect from him all such positive good offices and disinterested services as the moral improvement attained by mankind has rendered customary, he deserves moral blame if, without just cause, he disappoints that expectation. Through this principle the domain of moral duty, in an improving society, is always widening.¹²

It is, finally, interesting to note that a development of Hume’s account of our evaluation of agents is not only able to accommodate our common-sense intuitions about supererogatory actions, but can do so while avoiding some of the main problems faced by standard contemporary accounts of supererogation. According to contemporary act utilitarianism, any action other than the one that is optimific is morally wrong even if it displays a very unusual degree of altruism. For example, according to act utilitarianism even someone who donates a huge

proportion of their income to aid agencies is still acting wrongly if an even larger donation would promote more overall welfare. Its failure to allow for any supererogatory actions has cast serious doubt on the viability of the theory. By contrast, Hume holds that our evaluation of agents depends on what we can expect of them, knowing a “degree of selfishness . . . to be inseparable from human nature.” Heroic and saintly actions involve the agent’s sacrificing some of her own central interests. Performing such actions will therefore involve overcoming the very strong pull of self-interest, and so will be extremely difficult and unusual. On Hume’s account, this difficulty means that we cannot expect agents to act in this way, and excuses agents who fail to do so. But it would seem that the converse also naturally follows: that someone who does make such a sacrifice, despite the extreme difficulty of doing so, far exceeds our normal expectations of agents. This would account for our intense admiration for those who perform heroic or saintly actions, while retaining an account of the moral point of view as fully impartial.

This picture avoids many of the paradoxes and tensions faced by standard accounts of supererogation. Standard accounts hold that the reason heroic and saintly acts are beyond the call of duty is that the agent’s reasons of self-concern not to perform them have intrinsic moral weight, and so can counterbalance the impartial moral reasons to perform the action. This means that the balance of the moral reasons does not dictate that the agent perform the heroic or saintly action. In this way, these accounts of supererogation offer an explanation for why the agent is morally free not to perform that action. However, they then all face the problem of showing why performing the supererogatory act is more morally praiseworthy than acting on the reasons of self-concern, if the latter also have intrinsic moral weight and can balance the impartial moral reasons in favour of performing the action. On Hume’s picture, by contrast, the strong pull of self-concern that opposes performing the heroic action or saintly has no intrinsic moral weight, but is rather a moral limitation. It is therefore clear why it is most morally praiseworthy to perform the supererogatory action.

I conclude that Hume holds that the moral point of view is fully impartial, but that his position nevertheless accommodates some central common-sense intuitions concerning the impact of the force of self-interest on our expectations and moral evaluations of agents.

I will now turn to the question of where Hume’s moral theory stands with respect to the third tenet of utilitarianism.

3. Utility as a Single Measure of Interpersonal Value

The third tenet of utilitarianism Sayre-McCord identifies is “that utility be a single measure of interpersonal (though perhaps heterogeneous) value.” Sayre-McCord points out that this tenet “imposes two distinct but related constraints. That utility

constitutes a single measure is important for capturing the utilitarian commitment to the commensurability of consequences. That the value is interpersonal is important for distinguishing utilitarian views that appeal to the utility for all from egoistic and contractarian conceptions that rely only on the value for one or each (respectively)" (287).

Sayre-McCord's argument against the first constraint is that Hume's theory denies that utility serves as a single underlying criterion by which to solve all moral questions. His argument against the second constraint is that when Hume's account of morality is concerned with utility, it is primarily concerned with the interests of each, rather than the interests of all, and is closer to Hobbes's account than to Mill's. I will now consider these two arguments in turn.

3A. *Utility as a Single Underlying Criterion*

Sayre-McCord holds that Hume's moral theory avoids a commitment to a single overarching measure of moral value. He argues that the virtues serve an irreducible plurality of sometimes incompatible ends: "Virtues . . . are traits that . . . are well suited to the solving of certain problems posed by the circumstances and necessities of humankind," and "a trait may be well suited to the achieving of the relevant ends even if using it to secure those ends will frustrate others (as sometimes justice frustrates benevolence)" (282).

However several passages, such as the following one, indicate that the promotion of human interests constitutes, for Hume, a single underlying principle by which to decide solutions to various moral questions:

wherever disputes arise, either in philosophy or common life, concerning the bounds of duty, the question cannot, by any means, be decided with greater certainty, than by ascertaining . . . the true interests of mankind. (EPM 2.17: SBN 180)

Given Hume's many expressions of a commitment to welfarism, it is most plausible to interpret "the true interests of mankind" along straightforwardly welfarist lines.

The conflict between justice and benevolence which Sayre-McCord describes in support of his Bauhaus interpretation is discussed by Hume in several passages. One such example is when a law of justice prescribes taking money from a poor man and giving it to "a profligate debauchee," who "wou'd rather receive harm than benefit from large possessions" (T 3.2.1.13; SBN 482). Again Hume gives a straightforwardly welfarist account of how to resolve the conflict: we ought to act justly, since maintaining the institution of justice produces the most benefit overall.

3B. Utility as a Measure of Interpersonal Value

Sayre-McCord's argument against the view that Hume sees public utility as a measure of interpersonal value is that Hume's moral theory appeals to the interests of each rather than of all. In defense of this interpretation Sayre-McCord focuses on Hume's account of justice and argues that it is contractarian rather than utilitarian, and that it is closest to an egoistic Hobbesian account. I will first argue against an egoistic Hobbesian interpretation of Hume's account of morality, and will claim that Hume's account is not egoistic but impartial.

I will then turn to the question of whether this impartiality should be interpreted along utilitarian lines, as appealing to "the interests of all," or along contractualist lines, as appealing to "the interests of each." I will argue that Hume does take the general good to be an interpersonal value consisting in the interests of all, but that his account of the general good differs significantly from that of so-called "classical utilitarianism."

3B i. Against a Hobbesian Interpretation

Sayre-McCord's first argument is that both the origin and justification of justice are based on the self-interest of each of the participants: "When it comes to questions of justice, Hume explains both the origin of, and the obligation to, justice by appeal to the interests of each rather than the interests of all." Sayre-McCord secondly appeals to "the extent to which, according to Hume, justice is . . . bounded by mutual advantage" (290).

Sayre-McCord is right that on Hume's account the origin and the scope of justice are grounded on the mutual self-interest of the participants, and that in this respect it closely resembles Hobbes's account. Justice consists in a set of property rules that develop as a convention to deal with our limited generosity and the scarcity of goods. Our passions of self-interest and love of those we are close to, if unregulated, would be extremely destructive to society. Property rules develop as a device to regulate these passions. We can each see how we are benefited by there being established property rules, and because of this we are each reliably motivated to comply with them.

However, Hume draws a sharp distinction between the origin of justice, in our mutual self-interest, and its moral justification. After his discussion of the origin of the institution of justice, Hume says: "We come now to the second question we propos'd, viz. *Why we annex the idea of virtue of justice, and of vice to injustice.*" The way the moral sentiments enter the story is through the faculty of sympathy, through which we are impartially concerned with the prejudicial impact on human society of violations of the property laws:

when the injustice is so distant from us, as no way to affect our interest, it still displeases us; because we consider it as prejudicial to human society,

and pernicious to every one that approaches the person guilty of it. We partake of their uneasiness by *sympathy*; and as every thing, which gives uneasiness in human actions, upon the general survey is call'd *vice*, and whatever produces satisfaction, in the same manner, is denominated *virtue*; this is the reason why the sense of moral good and evil follows upon justice and injustice. (T 3.2.2.24; SBN 499)

Our moral approbation for justice comes to be “annexed” only once the system of property laws is in place, and we feel concern, through “sympathy . . . upon the general survey,” for the impact of violations of such laws on whoever is affected, regardless of whether or not the violation affects our own interests.

Justice is morally appraised, then, from the point of view of “sympathy . . . upon the general survey,” which, as I argued in the last section, is fully impartial. And from this impartial point of view, the extent to which the scope of justice is constrained by mutual advantage is viewed as a moral limitation and another regrettable concession to psychological realism. The virtue of justice is a pragmatic response to the “*selfishness and limited generosity*” of humankind (T 3.2.2.16; SBN 494). Hume says that if someone “unacquainted with human nature . . . deliberates with himself what rules of justice or property would best promote public interest. . . . His most obvious thought would be, . . . to assign the largest possessions to the most extensive virtue, and give every one the power of doing good, proportioned to his inclination” (EPM 3.23; SBN 193). This closely resembles the model of the benevolent ideal spectator. Hume says that another obvious thought would be to distribute property equally, given diminishing marginal utility: “wherever we depart from this equality, we rob the poor of more satisfaction than we add to the rich” (EPM 3.25; SBN 194). But these schemes both fail to accommodate the “returning . . . selfishness of men” (EPM 3.7; SBN 186), and for that reason are utopian and would in fact have “pernicious” consequences. While it can be argued that Hume’s endorsement of the established property laws makes too much concession to the selfishness and limited generosity of human nature, it is striking that he holds that the ideal property laws, were they to be feasible, would be strongly egalitarian.

The point of justice is to provide a remedy for the concurrence of our “selfishness and limited generosity” and the situation of moderate scarcity. We can be motivated to adopt and comply with the conventional property laws because they serve the self-interest of each of us. A scheme that distributed property more equally and required a greater degree of impartiality from its participants would be “impracticable” (EPM 3.26; SBN 194). The merit of justice derives from the fact that having a set of property laws is vital for social cohesion, and the existing laws have better consequences than the alternative options given that a more equitable system of property laws would require a greater degree of human impartiality than we can rely on.

The laws of justice, then, are based on human nature, which is mid-way between being completely selfish and completely humane:

Implant in the human breast perfect moderation and humanity, or perfect rapaciousness and malice: By rendering justice totally *useless*, you thereby totally destroy its essence, and suspend its obligation upon mankind. (EPM 3.12; SBN 188)

If we were less self-interested, the “cautious, jealous virtue” of justice would no longer be needed. It could be replaced by “much nobler virtues and more valuable blessings” (T 3.2.2.16; SBN 5).

As evidence for his Hobbesian reading of Hume, Sayre-McCord appeals to Hume’s discussion of animals, which, he says, illustrates “the extent to which, according to Hume, justice is generated and bounded by mutual advantage” (290). Since animals are powerless to retaliate and make their resentment felt, the strong do “not, properly speaking, lie under any restraint of justice with regard to them” (EPM 3.18; SBN 190). However, since Hume takes the very function of justice as being to accommodate the moral limitations inherent in human nature (together with the condition of moderate scarcity), it is a mistake to infer anything about Hume’s account of morality as a whole from his account of justice. The virtue of justice, for Hume, has a very limited purview. It is concerned only with property laws, and its function is to regulate our potentially destructive passions of self love and love of those close to us by institutionalizing deterrents to stealing property. This deterrent-based institution simply has no application when it comes to our dealings with animals. Unlike Hobbes, Hume does not infer from the fact that animals are powerless to retaliate that we are under no moral restraints towards them. Rather, virtues other than that of justice govern our treatment of them. Moreover, these virtues are, as we have seen, “far nobler” than that of justice precisely because they do not have to make such a concession to human selfishness.

There is an important contrast with Rawls, who also talks about the circumstances of justice, drawing on Hume. Like Hume, Rawls holds that the laws of justice are a convention the participants arrive at in part at least because they are mutually advantageous. Unlike Hume, however, Rawls takes the moral legitimacy of the laws of justice to come directly from that hypothetical contract. There is therefore no further perspective from which to assess whether the terms of reciprocity built into them are morally good or bad. (As Brian Barry argues, Rawls’s justification of the rules of justice in terms of mutual advantage is in conflict with his primary justification for the rules of justice, which is based on impartiality.¹³)

For Hume, by contrast, the institution of justice is morally appraised from an independent, fully impartial general point of view consisting in concern

for the interests of all. From this perspective, our self-interest, and the limits it imposes on the scope of justice, is viewed as a moral limitation. So, whereas for Rawls justice is the first virtue of social institutions, for Hume “the cautious, jealous virtue of justice” is a pragmatic, rather ignoble system of institutionalized deterrents.

I conclude, then, that Hume does not offer an egoistic, Hobbesian account of morality as a whole. While the institution of justice does indeed accommodate passions of self-interest and of love of those close to us, since its function is to regulate such passions, its moral justification is based on the vital role it plays in protecting the general good.

The question remains, however, of whether the general good should be interpreted along utilitarian lines as an interpersonal value consisting in the interests of all, or along contractualist lines, as consisting in the interests of each *seriatim*.

3B ii. Utility as a Measure of Interpersonal Value

I have argued that on Hume’s account, adopting the moral point of view consists in having an impartial concern for the interests of all, through the mechanism of sympathy. This account of the moral point of view would suggest an interpretation of the general good as being a measure of interpersonal value. Rawls in fact describes Hume as offering a prototype of a conception of the ideal observer that can provide a deductive basis for deriving classical utilitarianism:

Suppose that the ideal observer is thought of as a perfectly sympathetic being. Then there is a natural derivation of the classical principle of utility along the following lines. An institution is right, let us say, if an ideally sympathetic and impartial spectator would approve of it more strongly than any other institution feasible in the circumstances. For simplicity we may assume, as Hume sometimes does, that approval is a special kind of pleasure which arises more or less intensely in contemplating the workings of institutions and their consequences for the happiness of those engaged in them. This special pleasure is the result of sympathy. In Hume’s account it is quite literally a reproduction in our experience of the satisfactions and pleasures which we recognize to be felt by others. Thus an impartial spectator experiences this pleasure in contemplating the social system in proportion to the net sum of pleasure felt by those affected by it. . . . He imagines himself in the place of each person in turn, and when he has done this for everyone, the strength of his approval is determined by the balance of satisfactions to which he has sympathetically responded.¹⁴

Hume, though, does not express any commitment to the view that the best outcome is whatever instantiates the greatest net sum total of welfare. In fact, he assumes, overly optimistically, that the interests of all and of each will generally coincide, and therefore does not seriously address the problem of how to resolve conflicts between different persons' interests. Nevertheless a few passages in Hume indicate that a scheme of justice may not be acceptable to and agreed to by each individual, and show a commitment to interpersonal trade-offs of some kind. At EPM Appx. 3.3 (SBN 304), he describes the scheme of justice as "concurred in by the whole, or the greater part of society." Again, in the following passage he argues that it is not possible for a scheme of justice to avoid burdens to some and that it is sufficient for an institution to qualify as just if it promotes "the balance of good":

Public utility requires that property should be regulated by general inflexible rules; and though such rules are adopted as best serve the same end of public utility, it is impossible for them to prevent all particular hardships. . . . It is sufficient, if the whole plan or scheme be necessary to the support of civil society, and if the balance of good, in the main, does thereby preponderate much above that of evil. (EPM Appx. 3.6; SBN 305)

This passage concerning the balance of good is strikingly similar to Bentham's mature formulation of the principle of utility:

I recognize, as the all-comprehensive, and only right and proper end of Government, the greatest happiness of the members of the community in question: the greatest happiness of all of them, without exception, in so far as possible: the greatest happiness of the greatest number of them, on every occasion on which the nature of the case renders the provision of an equal quantity of happiness for every one of them impossible: it being rendered so, by its being a matter or necessity, to make sacrifice of a portion of the happiness of a few, to the greater happiness of the rest.¹⁵

Both passages indicate a concern that as few as possible lose out in the distribution of happiness. The ideal goal of the distribution of happiness is to ensure its universality over the whole community, and particular hardships are justifiable only when they are "impossible" to avoid. Moreover, the passage in Bentham above explicitly expresses a concern for an equal distribution of happiness: the goal is that "an equal quantity of happiness" be distributed as broadly as possible.

Bentham struggled with various formulations of the principle of utility, and was never entirely satisfied with them. Admittedly, he did eventually arrive at the formula of "the greatest happiness"; but, as Fred Rosen points out, he gave up his

original formula “the greatest happiness of the greatest number” in favour of “the greatest happiness” because of his concern that the original formula might appear to unsympathetic readers to justify imposing huge burdens on the minority whenever a majority stood to benefit.¹⁶

There is a tension between these passages in Bentham that express a concern for the distribution of happiness and other passages in which he simply advocates maximizing the net sum total of happiness. Moreover, his formula “the greatest happiness of the greatest number” notoriously contains co-dependent variables. These tensions and ambiguities were redressed by subsequent utilitarians, and the simpler formulation according to which the goal of the principle of utility is to promote the greatest net sum total of welfare was taken to be definitive of the theory. However, I will now briefly suggest that the passages which advocate a more complex account of the general good, together with the passage from Hume cited above, follow much more naturally from the first two tenets of utilitarianism, on the most plausible interpretation of these tenets. I will also suggest an alternative to Rawls’s way of interpreting Hume’s impartial sympathetic spectator, and will argue that it implies this more complex account of the general good. Of course, a full defense of this argument would require a much longer discussion, which is beyond the scope of this paper, so my argument will be unavoidably sketchy.

So-called “classical utilitarianism” relies on an impersonal conception of welfare, according to which welfare is viewed for its moral significance in abstraction from its instantiation in particular persons’ lives. However, it is more plausible to take the moral importance of welfare to be grounded on the moral importance of the persons whose welfare it is; welfare matters because it is the welfare of people (or sentient beings in general, but for simplicity I will focus on people) who experience it in their lives.¹⁷ On this view our concern for welfare is based on concern for persons: concern for persons is manifested in concern for their welfare, because welfare is, by its very nature, what is beneficial to someone.¹⁸ Likewise, it is more plausible to interpret the utilitarian conception of impartiality as grounded on the equal moral status of persons: each person’s welfare matters equally because each person matters equally. In short, promoting welfare matters because persons matter.

If concern for welfare is grounded on concern for the persons whose welfare it is, then this concern, I suggest, is directed at the way in which welfare is instantiated in persons’ lives. Generally, we are concerned with the welfare people attain across their whole lives, that is, their global welfare, but in certain circumstances (such as those involving extreme pain) we are concerned with persons’ immediate fates. I will refer to both persons’ global welfare and their immediate fates as “how persons fare.” Impartiality is accordingly taken to be the claim that it matters equally how each person fares. As I will now argue, concern for how persons fare cannot be captured by concern for the net sum total of welfare, since knowing this sum total tells us nothing about how any of the individuals fare.

The contrast between this conception of welfare as grounded on concern for persons and the impersonal conception that classical utilitarianism is based on can be starkly illustrated by considering the following two scenarios: in the first, millions of people each suffer agony for ten seconds each, and in the second, one person suffers agony for hundreds of thousands of years. Let us suppose that the first scenario instantiates a slightly greater net sum total of agony. (For simplicity, we can assume that each ten second period of agony is equally bad. We can also assume the persons each die as soon as the period of agony is over, in order to avoid complications about how various periods of agony would impact on the other prudential values in persons' lives; since each person's life contains nothing but agony, they are each benefited by dying as soon as possible.) Classical utilitarianism would take the first scenario to be worse, given that it instantiates a greater net sum total of pain. It evaluates the moral significance of pain independently of its instantiation in particular individuals' lives, and, hence, independently of the way in which it is divided up among the lives.

By contrast, if our concern for the suffering is grounded in concern for the persons undergoing it then the focus of our concern, I suggest, is what these persons go through. In other words, our concern is directed at how the various individuals fare, rather than at the total quantity of pain experience per se considered in abstraction from the way it is instantiated in particular persons' lives. From this moral perspective it matters greatly whether or not the pain experience is all instantiated in a single person's life. (Derek Parfit's Repugnant Conclusion can be seen as illustrating the converse problem with classical utilitarianism. He points out that it would judge a scenario in which a vast number of people lead lives barely worth living to be better than one in which a few billion people lead richly fulfilled lives, as long as the first scenario instantiates a greater sum total of welfare.¹⁹ The problem again, I suggest, is that classical utilitarianism is concerned only with the total quantity of welfare and gives no weight to the way in which that welfare is instantiated in persons' lives—in this case, to whether it is concentrated in a smaller number of richly fulfilled lives, or so thinly spread that the lives in which it is instantiated are barely worth living. Again, it gives no *intrinsic* weight to how persons fare, and may judge to be preferable a scenario in which no one attains a satisfactory level of global welfare.)

The pain example, I suggest, demonstrates the extent to which classical utilitarianism distorts the nature of sympathy that is at the heart of the theory: we naturally feel a far greater sympathetic concern at the second scenario than the first. And the reason for this, I suggest, is that our sympathetic concern for suffering is naturally directed at the persons experiencing it, rather than at the amount of pain experience in the universe abstracted from the way it is experienced by the individuals involved. (Likewise, in the Repugnant Conclusion case, we naturally feel much greater sympathetic pleasure at a scenario in which a few billion people

lead richly fulfilled lives than one in which trillions lead lives barely worth living, again because our sympathy is not naturally directed at pleasure experience considered independently of the way it is instantiated in persons' lives.)

The pain example also demonstrates the inadequacy of the classical utilitarian conception of impartiality. Classical utilitarianism gives equal moral weight to equal quantities of pain, however they are distributed among lives. In doing so, however, it assigns no weight to the way in which the pain is instantiated in the lives of particular individuals, to the point that it treats as irrelevant whether the pain is all experienced in a single person's life so that that person suffers an unimaginably awful fate or whether it is so thinly distributed that the fate experienced by each individual is relatively trivial. The only weight it gives to the way in which individuals participate in the distribution of welfare is as a tie-breaking principle: in cases in which the total quantity of welfare is identical in each outcome we ought to choose the outcome in which it is distributed more equally. But as the pain example illustrates, as soon as the total quantity of pain is slightly smaller in one outcome, that outcome is to be preferred no matter how great the burden this would impose on someone. In judging the second outcome to be preferable, classical utilitarianism cannot be plausibly said to be giving equal weight to the persons involved.

The pain example therefore illustrates H. L. A. Hart's objection that the principle of equality that classical utilitarianism relies on is formal and empty. As Hart puts it: "Individual persons and the *level* of an individual's happiness are for the utilitarian only of instrumental, not intrinsic importance. Persons are merely the 'receptacles' for the experiences that will increase or diminish aggregate welfare. So utilitarianism is 'no respecter of persons' in a sinister as well as a benign sense."²⁰ In the pain example, classical utilitarianism treats persons as "mere receptacles" for pain experience in the sense that it is concerned only with the conglomerate total quantity of pain there is in the universe, and assigns no moral significance to the way in which that pain is experienced by the individuals involved. That is, it gives no intrinsic weight to what the individuals go through. This indicates that in order to adequately capture the equal moral importance of persons we ought to give equal moral weight to how each person fares, rather than to equal quantities of pain regardless of how they are divided up among lives.

I suggest, then, that if we interpret welfarism and impartiality as grounded on the moral importance of the persons whose welfare it is, then the focus of our moral concern should be how persons fare. Impartiality should be interpreted as the claim that it matters equally how each person fares. The general good can no longer be viewed as the conglomerate net sum total of welfare, since knowing this sum total does not tell us about how any of the individuals fare, as was illustrated by the pain example and by the Repugnant Conclusion. In order to know how persons fare we need to know how they participate in the sum total. On this conception of welfare, therefore, concern for the way in which welfare is distributed

among the individuals involved is built into the moral point of view. And since it matters equally how each person fares, then equality is not simply a tie-breaking principle. Rather, it is at the foundation of the principle of utility itself: how each person fares is a distinct moral value of equal moral weight, which therefore has a separate and equal claim on our attention when we adopt the moral point of view. Adopting the moral point of view consists in having an equal concern for how all of the various individuals fare. The general good is viewed not as the conglomerate net sum total of welfare, but as the *inclusive combination* of each person's global welfare (or immediate fate).

This conception of the general good is, I suggest, expressed in the passages from Hume and Bentham cited above. Both passages express a concern for how individuals participate in the distribution of welfare, and take as the ideal goal that the benefits are universally shared by the individuals involved. The passage in Bentham explicitly states as the ideal goal that each person participates equally in the distribution of happiness. Furthermore, key passages in both Bentham and Mill express the utilitarian conception of impartiality as the claim that each person has equal claim to the means of happiness, which implies a conception of impartiality according to which equal moral weight is assigned to how each person fares over the course of their lives (that is, their global welfare).²¹

I will now finally suggest that Hume's account of the ideal sympathetic spectator, or at least a natural development of it, represents an account of the moral point of view that implies this conception of the general good. Rawls rightly points out that on Hume's account of the mechanism of sympathy, the pleasures and pains of others are reproduced in the ideal sympathetic spectator. However, Rawls's assumption that the pleasure Hume's ideal spectator feels is proportionate to the net sum total of pleasure follows only on the assumption that the spectator sympathizes directly with pleasures and pains, considered in abstraction from their instantiation in persons' lives, rather than with the individuals experiencing those pleasures and pains.

It is more plausible, I suggest, that Hume takes the ideal spectator's sympathy to be directed not at pleasures and pains taken independently, but rather at the individuals experiencing them. I have argued that on Hume's account, adopting the moral point of view involves correcting for the effect of contiguity and partiality on our felt passions, so that that we form the moral judgment that the welfare of those distant to us matters morally just as much even though the sympathy it elicits in us is "far from being as lively as when our own interest is concern'd, or that of our particular friends" (T. 3.3.1.18: SBN 583). This suggests that at the foundation of the moral point of view is concern that is focused on particular individuals—ourselves, and those we are close to. Adopting the moral point of view simply involves extending to everyone this same concern, although, given that it is impossible to have the same intensity of felt concern for distant strangers

as we have for those we are close to, we arrive at the moral point of view largely by extrapolating from our self-concern and our concern for those we are close to. The concern we feel from the moral point of view is therefore a calm passion.

I suggest, then, that on Hume's account of ideal sympathetic spectator, the spectator's concern is directed at the welfare of particular individuals. This suggests a picture of the moral point of view as being omnipersonal rather than impersonal: it consists in individualized concern for the welfare of all the various individuals. Moreover, since we arrive at the moral point of view by extrapolating from our concern for our own welfare (and that of those we are close to), the idealized concern we feel from the moral point of view for each individual's welfare is as strong as that person's self-concern (though it lacks the same felt intensity, and consists in a calm passion).

Since the spectator's sympathetic concern for pleasures and pains is directed at the individuals experiencing them, this will not plausibly lead to the "classical utilitarian" conception of the general good as the greatest conglomerate net sum total of pleasure or pain. As Rawls points out, the account of the ideal sympathetic spectator that grounds a rationale for choosing the greatest net sum total of happiness is one in which the distinct experiences of different individuals are amalgamated within a single individual's perspective. The spectator imagines that he himself were to experience all the pleasures and pains felt by the individuals involved. This explains the choice of the second scenario in the pain case, for example: if we imagine that we ourselves would experience all the pain experience, then the rational choice is whichever option instantiates the smallest sum total of pain experience. As Rawls argues, this ignores the moral significance of the fact that the experiences are in reality instantiated in distinct persons' lives. But if the spectator's sympathetic concern for the pleasures and pains is directed at the persons experiencing them, then, I suggest, the choice of the second scenario lacks a plausible rationale.²² On the contrary, the scenario in which the pain is all experienced in a single person's life evokes a far greater sympathetic pain than the first scenario.

Rather, as I have argued, if the spectator's concern for the pleasures and pains is grounded on concern for the persons experiencing them then this concern is directed at the way the pleasures and pains are instantiated in persons' lives (that is, at how persons fare). How each person fares is a distinct value, and knowing how the various individuals fare requires knowing how they each participate in the conglomerate sum total. In addition, how each person fares matters equally, and has an equal claim on our attention from the moral point of view.

Concern for one person is manifested in concern that that person fares as well as possible. On Hume's account of the sympathetic spectator, I have argued, adopting the moral point of view involves extending this concern to everyone equally. It is therefore manifested in concern that they all fare as well as possible, giving

equal weight to how they each fare. This is expressed in Bentham's account of the ideal goal as "the greatest happiness of all, without exception," and as "an equal quantity of happiness for every one of them." It is also expressed in the passage from Hume (cited in section one) in which Virtue describes as her sole purpose "to make her votaries and all mankind, during every instant of their existence, if possible, cheerful and happy" (EPM 9.15; SBN 279); her goal is for each person to attain happiness during the whole course of their lives (that is, for each person to attain a high level of global welfare).

This omnipersonal perspective of sympathetic concern for all is a synoptic perspective, which includes within it the moral importance of how everyone fares. However, on this interpretation, the distinct experiences of different individuals are not amalgamated within the spectator's single perspective in the sense of giving no intrinsic weight to the fact that the pleasures and pains are instantiated in the lives of distinct individuals. Moreover, since the spectator has a separate and equal concern for how each person fares that is as intense as that person's self-concern, he is concerned that no-one misses out in the distribution of happiness. Hardships to particular individuals are viewed as deeply regrettable, and are justified only when they are "impossible" to avoid. The ideal goal is for each person to enjoy a flourishing life.

Hume's account of the general good is very brief and offers no details of how this ideal goal might be implemented, but the phrase "public utility" at the beginning of the passage at the start of this section suggests that it is best implemented by promoting common goods that everyone has an interest in. This phrase strongly resembles Bentham's phrase "the universal interest," which he describes as the ultimate end of government.²³ Bentham develops this account at length, and discusses what he takes to be the most important universal interest, security, which includes the need for an adequate subsistence income as well as security against assault and theft.²⁴ Both he and Mill argue that security must be guaranteed to everyone.

Nevertheless, there may be an unavoidable conflict between different persons' basic interests, so that it is not possible for each person's basic interests to be secured. The omnipersonal moral point of view (consisting in individualized concern for all) will then justify certain interpersonal trade-offs, although since these are governed by the goal of giving equal consideration to how each person fares, they will not be the same as those that would be justified by classical utilitarianism.

The omnipersonal moral point of view of impartial concern for the interests of all leads to an account of the general good as consisting in the interests of all rather than the interests of each *seriatim*. Given that everyone's interests are included within the moral point of view, the greater the number who suffer a certain hardship the greater the spectator's sympathetic pain and the worse

the situation is taken to be. Therefore intrinsic moral weight is assigned to the number who will experience a certain benefit or burden. This means that where there is an unavoidable interpersonal conflict of interests, a burden to some individuals might be viewed as outweighed by a burden imposed on a greater number of others.

On the one hand, then, the spectator's individualized concern leads to the goal that no-one misses out in the distribution of welfare. Since the spectator has a separate and equal concern for each person that is as strong as that person's self-concern, when anyone experiences a serious hardships this will be viewed as deeply regrettable. On the other hand, the breadth of the moral perspective of concern for all means that when it is unavoidable that some persons will suffer serious hardships, the greater the number who experience a hardship the worse the situation is taken to be. The spectator's concern for each person must always be balanced by concern for all the others. The least bad situation might not be judged to be the one that is individually least unacceptable, on the ground that this might give insufficient weight to the interests of all the rest.

Where the ideal goal that each person's interests are equally served by the distribution of welfare cannot be achieved, it may be best approximated by choosing the situation in which as many as possible share in the distribution and as few as possible suffer a burden, rather than the situation in which the individually greatest burden is as small as possible. Thus, Bentham describes the second-best ideal as "the greatest happiness of the greatest number of them on every occasion on which the nature of the case renders the provision of an equal quantity of happiness to every one of them impossible." Hume likewise argues that while the goal is for benefits to be universally shared (public utility), where hardships are unavoidable "the balance of good" is to be chosen: "though such rules are adopted as best serve the same end of public utility, it is impossible for them to prevent all particular hardships. . . . It is sufficient, if the balance of good, in the main, does thereby preponderate much above that of evil."

By contrast, from the contractualist moral point of view, such interpersonal trade-offs are ruled out. Situations are only morally assessed from the perspectives of single individuals. There is no moral perspective from which to judge that a situation in which a large number of people suffer a certain hardship is intrinsically worse, given that no single individual suffers the combined hardship. Therefore the moral perspective of concern for each *seriatim* leads only to a one-by-one comparison between the burdens and benefits experienced by each of the various individuals. Conflicts of interests are resolved by choosing the situation that is individually least unacceptable, that is, the situation in which the individually greatest burden is as small as possible.²⁵

The plausibility of some kind of interpersonal balancing of interests from the synoptic perspective of concern for all can be illustrated with the counter-example

Parfit gives to contractualism. He compares two scenarios. In the first, ten people suffer great agony for fifty years. In the second, ten million people suffer agony just as great for fifty years minus a day.²⁶ A day of agony is an enormous burden, and so if we compare the two scenarios from the perspective of concern for each one by one the first scenario is taken to be worse. Intuitively, though, taking the first scenario to be worse gives insufficient weight to the interests of all the other persons who will suffer hugely in the second scenario. Hume's account of the sympathetic spectator, I suggest, offers a forceful account of the perspective we adopt when we make intuitive judgements about what option would best promote everyone's welfare: we have a separate sympathetic concern for how each person fares that is as intense as that person's self-concern and we extend this concern to all of the individuals involved, and decide which scenario gives us most sympathetic pleasure or least sympathetic pain. This account is much less precise than that of classical utilitarianism. There is no formula by which to decide how to evaluate situations and decide interpersonal trade-offs. However, it is, I suggest, considerably more plausible.

I conclude that Hume's account of the sympathetic spectator suggests a conception of the moral point of view that does consist in concern for the interests of all rather than the interests of each seriatim, and that it therefore does take the general good to be a measure of interpersonal value. His account of the general good differs from that of so-called "classical utilitarianism," which is concerned only with the conglomerate sum total of welfare and gives no weight to the way in which the various individuals participate in it. I have argued, however, that it follows much more naturally from the most plausible conception of welfare and impartiality as grounded on the moral importance of persons, and moreover the conception that Hume and Bentham, along with many subsequent utilitarians, had in mind. I suggest, then, that in examining the most plausible utilitarian conception of the general good we should go back to the very origins of utilitarianism, in Hume's ideal spectator theory.

Conclusion

I hope to have shown that Hume adheres to the first two tenets of utilitarianism, welfarism and impartiality, while avoiding many of the problems faced by overly technical modern formulations of the theory. In particular, I have argued that Hume's account of the evaluation of agents accommodates common-sense intuitions concerning supererogation, and that his account of the moral point of view is impartial without being utopian. In the case of the third tenet of utilitarianism, I have argued that while Hume's account of the general good differs from that of so-called "classical utilitarianism," it is in fact a better expression of the most plausible conception of the first two tenets.

NOTES

I am grateful to Simon Blackburn, Roger Crisp, David Lyons, Michael Ridge, Gerald Postema, John Skorupski, and the editors and two anonymous referees at *Hume Studies* for their very helpful comments. I am also grateful to the Boston University Ethics reading group for their discussion of a draft of this paper.

1 It is beyond the scope of this paper for me to justify my viewing Hume's account of morality as prescriptive as well as descriptive, but this position has been persuasively defended. See, for example, Robert Shaver's argument for this view in "Hume's Moral Theory?," *History of Philosophy Quarterly* 12 (1995): 317–31.

2 For the view that Hume is some kind of utilitarian, see for example, Roger Crisp, "Hume on Virtue, Utility, and Morality," in *Virtue Ethics: Old and New*, ed. Stephen Gardiner (Ithaca: Cornell University Press, 2005, forthcoming); J. Plamenatz, *The English Utilitarians* (Oxford: Oxford University Press, 1967); R. J. Glossop, "The Nature of Hume's Ethics," *Philosophy and Phenomenological Research* 27 (1967): 527–36, at 535–6; J. Harrison, *Hume's Theory of Justice* (Oxford: Clarendon Press, 1981), 33, 88. For the opposing view, see, for example, B. Wand, "Hume's Non-utilitarianism," *Ethics* 72 (1962): 193–6; Douglas Long, "Utility and the 'Utility Principle': Hume, Smith, Bentham," *Utilitas* 2 (1990): 12–39.

3 For example, Annette Baier, argues along these lines in *A Progress of Sentiments* (Cambridge: Harvard University Press, 1991).

4 "Hume and the Bauhaus Theory of Ethics," *Midwest Studies in Philosophy: Moral Concepts*, ed. P. A. French, T. E. Uehling, Jr., and H. K. Wettstein (Notre Dame: University of Notre Dame Press, 1996): 280–98. All references are to this article, unless otherwise specified.

5 Thomas Scanlon, "Contractualism and Utilitarianism," in *Utilitarianism and Beyond*, ed. A. Sen and B. Williams (Cambridge: Cambridge University Press, 1982), 110.

6 The passage from Scanlon I cited above, for example, continues: "then it would seem to follow that the basis of moral appraisal is the goal of maximizing the *sum* of individual well-being" (110).

7 This interpretation of Hume's ideal sympathetic spectator as representing a fully impartial perspective of concern for the interests of all is widely held in the literature. See for example Ronald R. Glossop, "Is Hume a 'Classical Utilitarian'?" *Hume Studies* 2 (1976): 1–16, and John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1972).

8 "Hume on Motivating Sentiments, the General Point of view, and the Inculcation of 'Morality,'" *Hume Studies* 22 (1994): 36–58.

9 "The Sceptic," in David Hume, *Essays and Treatises on Several Subjects, vol 1: Essays Moral, Political and Literary*, ed. T. H. Green and T. H. Grose; in *Hume, The Philosophical Works* (Aalen: Scientia Verlag, 1964: reprint of the 1882 London edition), 3: 231. The essay was first published in 1742.

10 Peter Singer, *The Expanding Circle: The Ethics and Sociobiology* (New York: Farrar, Straus and Giroux, 1981).

11 For a subtle discussion of this aspect of Mill, see John Skorupski, *John Stuart Mill* (London: Routledge, 1989), 318–25.

12 *Auguste Comte and Positivism*, in *Essays on Ethics, Religion and Society*, in vol. 10 of *The Collected Works of John Stuart Mill*, ed. J. M. Robson (London and Toronto: Routledge and University of Toronto Press, 1969), 338 (quoted in Skorupski, 321).

13 Brian Barry, *Justice as Impartiality* (Oxford: Oxford University Press 1995).

14 Rawls, *A Theory of Justice*, 185–6.

15 *Parliamentary Candidate's proposed Declaration of Principles: or say, A Test proposed for Parliamentary Candidates* (London 1831), 7.

16 See Bentham's "Article on Utilitarianism," in *Deontology, together with A Table of the Springs of Action and Article on Utilitarianism*, ed. Amnon Goldworth (Oxford: Clarendon Press, 1983), 309ff. Fred Rosen discusses this in his *Jeremy Bentham and Representative Democracy, A Study of the Constitutional Code* (Oxford: Clarendon Press, 1983), 200–3.

17 In this discussion I will follow the classical utilitarian hedonistic conception of welfare, but my argument is intended to apply also to conceptions of welfare that include a non-experiential component.

18 In an extremely interesting and rich article, "Bentham's Equality-Sensitive Utilitarianism" (*Utilitas* 10 [1998]: 144–58), Gerald Postema discusses this conception of welfare, which he terms "the individualist conception," and argues that it can be attributed to Bentham. My discussion in this section is indebted to his article.

19 Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 387–90.

20 H. L. A. Hart, *Essays on Bentham* (Oxford: Clarendon Press, 1982), 99.

21 *The Works of Jeremy Bentham*, ed. J. Bowring (Edinburgh, 1838–1843; reprinted New York, 1962), 11 vols. 9: 108; J. S. Mill, *Utilitarianism*, chap. 5, para. 36 in *Collected Works*, ed. Robson, vol. 10.

22 It should be noted that the two main rationales for so-called "classical utilitarianism" that have been offered in recent years are both based on assigning no moral significance to the separateness of persons. Rawls argues that the sympathetic spectator that represents the classical utilitarian moral point of view falsely conflates the interpersonal and intrapersonal perspectives. Parfit suggests another possible rationale, grounded on the reductive account of personal identity. As he points out, if we adopt this reductive account, then personal identity and the distributive principles it supports may be assigned less weight. If personal identity is given no weight, this may support classical utilitarianism. According to this rationale, then, rather than falsely conflating the interpersonal case with the intrapersonal case, classical utilitarianism holds that whether or not the various instantiations of pleasure and pain occur within a single person's life is irrelevant in the intrapersonal as well as the interpersonal cases. See *Reasons and Persons*, 329–45.

23 *The Works of Jeremy Bentham*, 3: 452.

24 Gerald Postema discusses Bentham's account of the universal interest, and the way in which it follows from his account of welfare as grounded on the moral importance of the persons whose welfare it is, in "Interests, Universal and Particular: Bentham's Utilitarian Theory of Value," *Utilitas*, forthcoming.

25 It should be noted that on Scanlon's account of Kantian contractualism, the magnitude of a burden is a function not just of a person's absolute level of well-being under general adoption of a certain principle, but also of how much that person could have been benefited by alternative principles. This means that the person with the individually greatest complaint is not necessarily the worst-off person, if, for example, the individual who is worst-off under a certain principle could only have been benefited very little by the adoption of alternative principles, whereas another individual who is better off but still has a low absolute level of well-being could have been benefited considerably by an alternative principle.

26 *Reasons and Persons*, 393.