The Appeal of Utilitarianism

ROBERT SHAVER
University of Manitoba

Utilitarianism continues to vex its critics even in the absence of generally respected arguments in its favour. I suggest that utilitarianism survives largely because of its welfarism. This explains why it survives without the backing of respected arguments. It survives without such arguments because justifying the value of welfare requires no such argument.

It is odd that utilitarianism continues to vex its critics even in the absence of generally respected arguments in its favour. For example, Philippa Foot records that 'utilitarianism tends to haunt even those of us who will not believe in it. It is as if we for ever feel that it must be right, although we insist that it is wrong.' Utilitarianism casts a 'spell' for which 'exorcism' is needed.[1] Judith Jarvis Thomson prefers a horticultural metaphor: utilitarianism 'keeps on reappearing, every spring, like a weed with a long root'.[2] Christine Korsgaard sees utilitarianism as a trap: 'We seem to succeed in disproving one utilitarian doctrine, only to find ourselves caught in the grip of another.'[3]

Foot, Thomson and Korsgaard would agree that utilitarianism does not owe its supernatural, weed-like, or gripping quality to the strength of the standard arguments made on its behalf. For these are unimpressive. Bentham seems to infer utilitarianism from the unattractiveness of 'asceticism' or 'caprice'. Mill's 'proof' is debated as an instance of obvious fallacies, or as a mistaken instance of these fallacies. No one takes it as a successful argument for utilitarianism. Sidgwick beats his critics to the punch and ends the *Methods* in despair. Nonetheless, utilitarianism survives.

Foot, Thomson and Korsgaard believe it survives because of its consequentialism. They set aside, as unimportant, the traditional utilitarian view about which consequences matter, namely that welfare and nothing else matters. They suppose utilitarians are quite willing to abandon this; Thomson even supposes that the 'sophisticated modern

[1] Philippa Foot, 'Utilitarianism and the Virtues', *Mind* 94 (1985), p. 196.

[2] Judith Jarvis Thomson, 'Goodness and Utilitarianism', *Proceedings and Addresses of the A. P. A.*, lxvii (1994), 7.

[3] Christine Korsgaard, 'The Reasons We Can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values', *Social Philosophy and Policy*, x (1993), 24.

© 2004 Cambridge University Press
Utilitas Vol. 16, No. 3, November 2004
DOI: 10.1017/S0953820804001153
Printed in the United Kingdom

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

utilitarian . . . may have no theory of the good at all'.[4] They follow Samuel Scheffler:

> utilitarianism refuses to fade from the scene in large part because, as the most familiar consequentialist theory, it is the major recognized normative theory incorporating the deeply plausible-sounding feature that one may always do what would lead to the best available outcome overall. Despite all of utilitarianism's faults (including, no doubt, its misidentification of the best outcomes), its incorporation of this one plausible feature is in my opinion responsible for its persistence.[5]

Foot, Thomson, Korsgaard and Scheffler see the utilitarian theory of the good as unimportant because they concentrate, to varying degrees, on an apparent oddity of deontological theories: The deontologist refuses to approve of committing one act, bad by his lights, to prevent the occurrence of more than one of the same bad acts. Consequentialist theories – whether they value only welfare or things in addition to welfare – do not face this paradox, and so they are, in this respect, attractive.

Avoiding the paradox does not, however, fully explain the attraction of utilitarianism. Utilitarianism has proved uniquely attractive among consequentialist theories. Even some non-consequentialists who suppose they have resolved the paradox continue to see an attraction.[6] And what pulls people towards utilitarianism in the first place is surely much simpler than paradox-avoidance: Utilitarianism is,

[4] Thomson, 'Goodness and Utilitarianism', 7. Korsgaard thinks that utilitarianism survives because it assumes that 'the business of morality is to *bring something about*' (Korsgaard, 'Reasons', 24). See also Foot, 'Utilitarianism', 197. Elsewhere, Foot attributes some additional appeal to 'the fact that utilitarians insist that moral goodness must somehow be connected with what is good to and for human beings'. See 'Morality, Action and Outcome', *Morality and Objectivity*, ed. Ted Honderich (London, 1985), p. 29.

[5] Samuel Scheffler, *The Rejection of Consequentialism* (Oxford, 1982), p. 4.

[6] Foot admits that utilitarianism may seem attractive because '[i]t cannot be enough to say that we *do* have such things as rules of justice in our present system of virtues: the question is whether we should have them, and if so why we should'. She goes on to write that '[i]n its most persuasive form' this thought 'involves a picture of morality as a rational device developed to serve certain purposes'. Her objection is that this picture of morality is 'a consequentialist assumption' one should reject. ('Utilitarianism', pp. 208–9) But her initial worry – that we should ask whether we should have a certain system of virtues – survives this rejection. For one need not think of morality as a device to have the worry that some particular moral code is not one we should have, on the ground that it is bad for us. Even if morality, properly understood, has no extra-moral purpose, one can worry that obeying it frustrates too many of our extra-moral purposes. Thus Korsgaard, who shares Foot's dislike of picturing morality as a device, is very exercised by the worry that morality is bad for us. (For the worry, see Christine Korsgaard, *The Sources of Normativity* (Cambridge, 1996), pp. 50, 61, 76. For the dislike of picturing morality as a device, see 'Reasons', pp. 25, 49–51.) Benevolence, understood as the virtue concerned with decreasing what is bad for us, seems to have a special role in judging the virtues.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

as J. J. C. Smart and John Harsanyi note, the theory that appeals to our benevolence.[7]

I suggest utilitarianism survives largely because it values welfare and nothing else.[8] This explains why it survives even without the backing of respected arguments. It survives without such argument because justifying the value of welfare requires no such argument. This may also explain why it has seemed so natural to make any ideal observer benevolent and why many believe that facts about welfare must be a necessary part of the content of morality.

Three preliminaries:

First, I take welfare to be the concept of which hedonism, preference satisfaction, or perfectionism are the usual specifications.

Second, I follow Amartya Sen, apart from substituting 'welfare' for 'utilities': Welfarism is the view that '[t]he judgment of the relative goodness of alternative states of affairs must be based exclusively on, and taken as an increasing function of, the respective collections of individual utilities in these states'.[9]

Third, to avoid generalities, let the anti-welfarist position be represented by Ross. I choose Ross on the basis of famous examples such as these:

> If we compare two imaginary states of the universe, alike in the total amounts of virtue and vice and of pleasure and pain present in the two, but in one of which the virtuous were all happy and the vicious miserable, while in the other the virtuous were miserable and the vicious happy, very few people would hesitate to say that the first was a much better state of the universe than the second.[10]

Ross thinks many would choose the first even if the total pleasure were greater in the second: '[S]uppose that A is a very good and B a very bad man, should I then . . . think it self-evidently right to produce 1,001 units of good for B rather than 1,000 for A? Surely not.'[11] Similarly,

> [s]uppose . . . that the fulfilment of a promise to A would produce 1,000 units of good for him, but that by doing some other act I could produce 1,001 units of

[7] J. J. C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge, 1973), pp. 4–8, 67; John Harsanyi, 'Morality and the Theory of Rational Conduct', *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams (Cambridge, 1982), pp. 56, 62; 'Problems with Act-Utilitarianism and with Malevolent Preferences', *Hare and Critics*, ed. D. Seanor and N. Fotion (Oxford, 1988), p. 96.

[8] For the same view, see T. M. Scanlon, 'Contractualism and Utilitarianism', in Sen and Williams, *Utilitarianism*, p. 108.

[9] See Amartya Sen, 'Utilitarianism and Welfarism', *Journal of Philosophy* 76 (1979), p. 468.

[10] W. D. Ross, *The Right and the Good* (Indianapolis, 1988), p. 138; see also pp. 26 f., 58, 136 f.

[11] Ibid., p. 35.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

good for B, to whom I have made no promise, the other consequences of the two acts being of equal value; should we really think it self-evident that it was our duty to do the second act and not the first?

Ross concludes: 'I think not.'[12]

In I, I give the argument for thinking welfare has a special status. In II, I consider objections. In III, I suggest one explanation for why welfare (at least on some construals) has this status.

I

In the debate between Ross and the welfarist, there is agreement that welfare is good.[13] No one, as Sidgwick notes, wants 'the paradoxical position of rejecting happiness as absolutely valueless'.[14] As Sidgwick argues, those, like Ross, who care about distribution or promise-keeping for its own sake, must value something else. Otherwise, they have nothing to distribute that, by their own lights, matters, and no plausible explanation of when promises are to be kept.[15] Ross can be described as adding something valuable – distribution, or promise-keeping – to what all agree has value – welfare. The characteristic welfarist thought is that this addition requires defence, in a way that the initial claim that welfare has value does not.

The addition requires special defence for two reasons. First, it is a contested addition. All agree that welfare is valuable; not all agree that Ross's promise should be kept, or with retributive punishment. Second, it is a suspect addition. It requires the reduction of something Ross, like everyone else, finds valuable.[16] This is the force behind the familiar worry that retributivists are simply malicious.[17] The worry

[12] Ibid., pp. 34–5.

[13] This section is taken from my 'Reflection and Well-Being', *Persons and Passions*, ed. Joyce Jenkins, Jennifer Whiting and Christopher Williams (Notre Dame, forthcoming).

[14] Henry Sidgwick, *The Methods of Ethics* (Indianapolis, 1981), p. 406. See also Francis Hutcheson, *An Inquiry concerning the Original of our Ideas of Virtue or Moral Good*, in *British Moralists*, ed. L. A. Selby-Bigge (Indianapolis, 1964), v. i, para. 112.

[15] Sidgwick, *Methods of Ethics*, pp. 392–3.

[16] For a similar point, see Shelly Kagan, *The Limits of Morality* (Oxford, 1989), pp. 16–18. The addition may be suspect for a further reason. F. C. Sharp argues that retributivism, in particular, does not cohere with the rest of our impartial, benevolence-directed moral practices, since it ignores the interest of the victim of the retribution. It is also parasitic on these other practices, since 'it demands retaliation only when the person to be punished has acted in violation of eudemonic standards'. See F. C. Sharp, 'Hume's Ethical Theory and its Critics II', *Mind* 30 (1921), pp. 154 f. and *Ethics* (New York, 1928), p. 137.

[17] Smart writes that '[t]here is *prima facie* a necessity for the deontologist to defend himself against the charge of heartlessness, in his apparently preferring abstract conformity to a rule to the prevention of avoidable human suffering' (Smart and Williams, *Utilitarianism*, p. 6; also pp. 5, 67, 72).

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

need not be a serious suspicion that malice is what really motivates Ross. It is rather the suspicion that the justification he can offer is really no better than the justification a malicious person might offer. Thus Sidgwick fears that the retributivist wants a 'purely useless evil', inconsistent with benevolence.[18] Of course nothing here shows that Ross *cannot* offer a better justification. The point is rather that he needs to give some justification, in a way that someone who thinks that welfare matters need not.[19]

At this point, the welfarist becomes sceptical. He is sceptical that much can be said in favour of promise-keeping or retribution. (The paucity of argument is particularly obvious in Ross, with his repeated appeals to what we see, or do not see, to be self-evident.) True, the welfarist may be unable to provide a defence of valuing welfare that goes further. If both Ross and the welfarist are faced with the need to justify what they value, they might well be driven to the same unhappy claim: 'We just do value this.' The argument turns on preventing this interrogation of both parties. One party, the argument claims, is a suspect; the other is not.

It is worth noting that Bentham may intend this argument. He sees many opponents of utilitarianism as 'partizan[s] of the principle of asceticism', 'approving of actions in so far as they tend to diminish . . . happiness; disapproving of them in as far as they tend to augment it'.[20] This strikes the reader as misguided. Opponents of utilitarianism need not be ascetics, and asceticism is not a popular position. But Bentham's point may be that opponents of utilitarianism (or at least welfarism) have no more justification for their position than

[18] Sidgwick, *Methods of Ethics*, p. 72. For the classic survey of rationales for retributivism, see J. L. Mackie, 'Morality and the Retributive Emotions', in Mackie, *Persons and Values* (Oxford, 1985).

[19] This thought has direct application to a recent defence of egalitarianism. Larry Temkin attacks 'the Slogan': 'One situation cannot be worse (or better) than another in any respect if there is no one for whom it is worse (or better) in any respect' (Larry Temkin, *Inequality* (Oxford, 1993), p. 256, emphases omitted). Temkin charges that anti-egalitarians simply cite the Slogan, without realizing that many people value things other than welfare. But one might see appeals to the Slogan in a more sympathetic light. The anti-egalitarian does not *simply* note that egalitarianism violates the Slogan. Instead, she notes that egalitarianism violates the Slogan and that, like everyone else, the egalitarian values welfare. She then asks what argument the egalitarian has in favour of equality such that reducing welfare is sometimes justified. The Slogan operates as a check on proposed outcomes. Where a proposed outcome promises a drop, offset by no gains, in welfare, a special justification is needed. Such a check seems quite appropriate, given that all too often we have found welfare-independent justifications to be bogus. Consider what we now make of arguments condemning homosexuality as unnatural or divinely proscribed, or of arguments approving of the subjection of women as natural or divinely ordained, or of arguments in favour of going to war so that national honour is recovered or defended.

[20] Jeremy Bentham, *Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart (London, 1970), *The Collected Works of Jeremy Bentham*, secs. II.IV, III, emphases omitted.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

ascetics do. Of course an opponent of welfarism can suggest principles, other than asceticism, that differ from utility. But here Bentham sees 'caprice' – the 'mere averment of . . . unfounded sentiments', foisted on us in a 'despotical' fashion.[21] Bentham's worry may be, again, that attempted justifications of principles other than utility will fail. Here the reader quickly asks whether Bentham has anything more impressive to say in favour of utility, and is disappointed to find that he does not. Caprice, it seems, is everywhere. But Bentham might reply that utility needs no defence: he claims that a proof of the principle of utility is 'needless', perhaps because there has never been anyone 'who has not on many, perhaps on most occasions of his life, deferred to it'.[22] Bentham can mean, as he says, that a proof of the principle of utility is 'impossible', without thus putting utility and its rivals on a par. The rivals need a proof; utility, because not challenged, does not.

II

One might object that scepticism about additions to welfare is a welfarist prejudice. The worry is that any scepticism is driven by a prior conviction that welfare alone matters, rather than by doubts specific to (say) the retributivist or egalitarian proposal. Ross may just appeal to intuition, and so is open to the charge of caprice. But those, like Kant, who share Ross's enthusiasm for retribution and promise-keeping, without sharing his intuitionism, have a great deal more to say. Whether this objection is persuasive depends on *what* more there is to say. Welfarists can, however, derive some confidence from noting that philosophers are much better at criticism than construction.

I want to concentrate on five more manageable objections.[23] The first finds my strategy suicidal. The next three query the claim that all agree that welfare is good by pointing to disagreements over theories of welfare, by noting examples of welfare that some do not find good, and by presenting philosophers who seem not to value welfare at all. The fifth objection claims that agreement has no epistemic force, and so, even if there is agreement on welfare and not on other goods, this does not show that welfare is privileged.

[21] See ibid., secs. I.XIV, II.XI–XIV, XIVn. R. M. Hare's treatment of 'ideals' as mere preferences reflects the same worry. See, for example, the exchange between Richard Brandt, Peter Singer and Hare in Seanor and Fotion, *Hare and Critics*, pp. 149–52, 220–1, 269–70. For an earlier version of the same thought, see Sharp, 'Hume's Ethical Theory', 155, and *Ethics*, p. 502.

[22] Bentham, secs. I.XI, XII; also I.XIII, II.VIII, X.

[23] For a defence of welfarism against very different objections, see Andrew Moore and Roger Crisp, 'Welfarism in Moral Theory', *Australasian Journal of Philosophy* 74 (1996).

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

1. My strategy requires that arguments for additions to welfare be shown to fail. Some of these arguments – such as those of Ross – rely on an appeal to self-evidence. The case for welfare also depends on an appeal to self-evidence. Hence, one might object, the case against Ross will also sweep away the case for welfare.

This is mistaken. The argument against Ross need not turn on a wholesale rejection of his intuitionism. It may turn, say, on rejecting the appeal to *contested* claims to self-evidence; the attraction of welfare is that its appeal is uncontested. Alternatively, the case for welfare need not appeal to self-evidence; one might find welfare valuable without thinking that anything is self-evident. Indeed, because the value of welfare is uncontested, *no* case for it needs to be made.

2. There is disagreement over the best theory of welfare. Some favour preference hedonism; others favour perfectionism or unrestricted preference satisfaction. Given this, one might object that agreement on the value of welfare is illusory.

This is unconvincing. Opponents in the debate over the best theory of welfare describe themselves as giving rival theories of one thing – welfare. There seems no reason to reject this description. For there certainly seems to be a pre-analytic concept – welfare, happiness, one's life going well – about which we have and share intuitions. Most accept various truisms about this concept: For example, welfare can increase or decrease; it is prima facie a good thing when welfare increases; prudence consists in increasing one's own welfare; benevolence consists in increasing the welfare of others; some non-human animals, such as my dog, have a welfare; evaluations of an individual's welfare are distinct from other evaluations, such as aesthetic or moral evaluations.[24]

3. The most popular worry about welfarism is that many do not value welfare derived from certain sources – from sadism or malice, for instance.[25] This jeopardizes the claim that all agree that welfare is good.

I want, first, to set aside a popular welfarist defence, offered most recently by L. W. Sumner. He writes that the solution lies

> in reminding ourselves how little welfarism determines of the structure of a moral theory: while it tells us wherein the good consists, it does not dictate how

[24] See L. W. Sumner, *Welfare, Happiness, and Ethics* (Oxford, 1996), ch. 1.

[25] See, for example, Sen, 'Utilitarianism', pp. 473–4, 477–8; Temkin, *Inequality*, p. 268; Ross, *The Right and the Good*, pp. 136–8; G. E. Moore, *Principia Ethica* (Cambridge, 1903), pp. 209–10, 213–14; John Rawls, *A Theory of Justice* (Cambridge, MA, 1971), pp. 30–1; C. D. Broad, *Five Types of Ethical Theory* (London, 1930), p. 234.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

> the right is to be derived from it. One possibility . . . is . . . utilitarian, where all welfare gains and losses are counted . . . . [But] welfarism leaves other directions open as well, including the option of ruling the deliberate infliction of harm on others impermissible in principle, whatever its overall welfare payoff. An ethical framework with this deontological structure is still welfarist, since it admits nothing but welfare to its theory of the good. It simply addresses the moral problem raised by evil appetites elsewhere, in its account of the right.[26]

This is unsatisfactory.[27]

The problem is that most critics reject the welfarist ranking of states of affairs, not merely utilitarian judgements about the right or about what should be done.[28] Suppose, as I am standing in a steamy and packed bus, I pass a smashed-up BMW and its unhappy owner. I feel pleased that the owner is unhappy. Stipulate that this feeling will not lead me to do anything nasty to other BMWs or their owners, nor will it make me less ready to assist them should the chance arise. Barring (as Sumner does) an implausibly moralized view of welfare, the welfarist will then judge that it is better that I am pleased by seeing the accident than that I am indifferent or sad about it. He will also judge that things improve the more pleasure I feel. Perhaps, as Sumner suggests, the welfarist can refuse to let this pleasure count in determining the right. But the critics of welfarism need not be objecting to any claim about the right. They are objecting to the welfarist's ranking of the states of affairs.[29]

Two replies do better. The first reply casts doubt on the anti-welfarist intuition when, as suggested, it concerns the goodness of states of affairs. Begin by considering how I might be criticized for feeling

[26] Sumner, *Welfare, Happiness, and Ethics*, pp. 199–200; also pp. 187, 218. For a similar solution, see James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford, 1986), pp. 25–6, 320–1 n. 13. Jan Narveson and Harsanyi argue that even utilitarians can ignore nasty preferences by the same separation of the right and the good. See Narveson, *Morality and Utility* (Baltimore, 1967), pp. 162–9; Harsanyi, 'Morality', 56, 'Problems', p. 96. For good criticism of utilitarian use of this strategy, see Madison Powers, 'Repugnant Desires and the Two-Tier Conception of Utility', *Utilitas* 6 (1994), pp. 171–6.

[27] In addition to the objection below, there is a further problem with Sumner's version of the defence. Welfarism is not simply the theory that 'admits nothing but welfare to its theory of the good'. As Sumner stresses elsewhere, the welfarist holds that 'welfare is the only value which an ethical theory need take seriously, ultimately and for its own sake' (Sumner, *Welfare, Happiness, and Ethics*, p. 3; also pp. 25, 184, 191). Welfarists 'affirm the priority of the good'. For them, 'the ultimate point of ethics is to bring about intrinsically valuable states of affairs' (p. 185). If so, it is implausible to think that a welfarist could rule out an action 'whatever its overall welfare payoff'.

[28] This is clear for Sen, Broad, Ross, Moore, and Temkin.

[29] This undercuts one of Hare's replies. Hare notes that the utilitarian will not recommend bear-baiting, because there are better alternative forms of entertainment (Hare, *Moral Thinking* (Oxford, 1981), p. 142; also Seanor and Fotion, *Hare and Critics*, pp. 246–7). But the utilitarian will prefer a pleased bear-baiting audience to an indifferent or miserable one, provided that these attitudes to entertainment are largely unconnected with what members of the audience are willing to do.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

pleased at the sight of the BMW, without the stipulation given above. One might say that my attitude is likely to result in my failing to help others, or even in my doing damage to other BMWs. One might say that, on these grounds, I am a poor person, or at least that I could have a better attitude – one more likely to lead me to help others. These criticisms are familiar from welfarists such as Smart, R. M. Hare and Yew-Kwang Ng.[30] The familiar objection is that this does not fully explain our criticism. Sen, replying to Hare and Ng, notes that

> it is certainly plausible that . . . most types of welfarists would find good reasons for developing a general 'anti' attitude [to me]. But that is quite a different claim from maintaining that the only reasons for the existence of these 'anti' attitudes are welfarist ones.[31]

C. D. Broad, attacking Sidgwick, admits that

> [n]o doubt malice is a state of mind which on the whole tends to increase human misery. But surely it is clear that we do not regard it as evil, simply as a means. Even if we were quite sure that all malice would be impotent, it seems clear to me that we should condemn it as intrinsically bad.[32]

Geoffrey Scarre, replying to Smart, writes that

> [s]adism and *Schadenfreude* have no place in our usual notions of legitimate pleasure sources. . . . If a soldier is under orders to kill his enemy in battle, we prefer him to kill with distate rather than relish. This is not because (or not primarily because) we fear that someone with a lust for killing is a danger to society. Our revulsion springs from a deeper level than that, a horror that such things can please people and a sense that whoever feels as he does falls well short of the human ideal. . . . [P]eople condemn the sadistic mentality not merely because sadists do damage but because wanting others to suffer is deplorable in itself.[33]

Now, it is hard to assess the welfarist claim, and anti-welfarist denial of the claim, that our attitudes are fully explained by welfarist thoughts. Even in cases where indirect welfarist considerations are excluded by stipulation, the welfarist can still say that we cannot escape our welfare-induced intuitions. I think the better route is to grant and then question the non-welfarist intuition.[34] When, on non-welfarist grounds, Broad sees clearly that malice is intrinsically bad, or Scarre finds himself revolted and horrified, judging that the human ideal is

[30] Smart and Williams, *Utilitarianism*, pp. 25–6; Hare, *Moral Thinking*, p. 141, *Hare and Critics*, p. 247; Yew-Kwang Ng, 'Welfarism: A Defence Against Sen's Attack', *Economic Journal* 91 (1981), pp. 527–30.

[31] Sen, 'A Reply to "Welfarism: A Defence Against Sen's Attack"', *Economic Journal*, xci (1981), p. 533; also 'Utilitarianism', pp. 476–8, 485–6.

[32] Broad, *Five Types of Ethical Theory*, p. 234.

[33] Geoffrey Scarre, 'Utilitarianism and Self-Respect', *Utilitas* 4 (1992), pp. 29, 32.

[34] Hare briefly notes this strategy: see *Moral Thinking*, pp. 142–3, 146. See also Peter Vallentyne, 'The Problem of Unauthorized Welfare', *Nous* 25 (1991), esp. pp. 298, 319 n. 5.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

not attained, one should ask what the non-welfarist grounds are, and how they differ from the grounds of those who see or feel that homosexuality is intrinsically bad or revolting. The obvious difference – that homosexuality is typically harmless whereas sadism is typically harmful – has been put aside. Again, a justification is particularly pressing given that Broad and Scarre presumably value other increases in harmless welfare.

The second reply makes more precise that to which there is agreement by distinguishing between final and unconditional goods.[35] Welfare is a final good if and only if it is good as an end – that is, it is not good merely instrumentally. Welfare is unconditionally good if and only if it is good in all circumstances. To illustrate the distinction, consider Kant. Kant holds that welfare is a final good. If it were not, happiness would not be a suitable reward for virtue, contrary to the doctrine of the highest good. Nor would benevolence be the virtue Kant believes it to be.[36] But, like those who object to welfare derived from sadism or malice, he does not hold that welfare is unconditionally good. Happiness is 'not of itself absolutely good in every respect but always presupposes conduct in accordance with the moral law as its condition'.[37]

Given this distinction, the argument of I can be recast: All agree that welfare is a final good; other candidates for being final goods are contested, and suspect when they require the reduction of what all take to be a final good. Of course, this would be less interesting were there agreement on some unconditional good. But there is no such agreement. (One might object that welfare is no longer unique: There will be widespread agreement on many goods as final goods, given the right conditions. But given the high number of welfarists – who value only welfare as a final good, whatever the conditions – the agreement cannot be too widespread.)[38]

[35] For the distinction, see Christine Korsgaard, 'Two Distinctions in Goodness', *Philosophical Review* 92 (1983), pp. 169–95.

[36] Ross makes both these points to show that 'the most austere moralists and the most anti-hedonistic philosophers are apt to betray the conviction that pleasure is good in itself' (Ross, *The Right and the Good*, p. 135).

[37] Immanuel Kant, *Critique of Practical Reason*, trans. Lewis White Beck (Indianapolis, 1956), Ak. 111, quoted by Ross, *The Right and the Good*, p. 136. See also Kant, *Foundations of the Metaphysics of Morals*, trans. Lewis White Beck (Indianapolis, 1969), Ak. 393, quoted by Temkin, *Inequality*, p. 278 n. 52.

[38] It is worth noting that Ross agrees, and thinks others agree, that, as I put it, welfare is a final good: '[W]e can all agree that the fact that a sentient being is in a state of pleasure is always in itself good, and the fact that a sentient being is in a state of pain always in itself bad, when this fact is not an element in a more complex fact having some other characteristic relevant to goodness or badness. And where considerations of desert or of moral good or evil do not enter, i.e. in the case of animals, the fact that a sentient being is feeling pleasure or pain is the whole fact (or the fact sufficiently described to enable us to judge of its goodness or badness), and we need not hesitate to say that the pleasure of animals is always good, and the pain of animals is always bad, in itself and

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

4. Bentham attacks ascetics. Say an ascetic is one who denies that welfare is a final good. Like Bentham, I do not think the existence of such people should be granted too readily. There are many people who recommend the sacrifice of welfare – their own or that of others – for, say, art or knowledge, but they see this as a sacrifice (even if an easy one). It is much harder to find someone who puts no value on welfare. One would place either no value or only instrumental value on involuntary sensual pleasures such as the warmth from a fire on a cold day. One would place no disvalue or only instrumental disvalue on pains such as a headache. One would have the same attitude to the welfare of others, which makes love and friendship impossible. True, many religious moralists discount sensual pleasures – but they are keen on stressing both the pleasures of heaven and the (usually sensual) pains of hell. Within philosophy, perfectionism is the obvious place to look for ascetics. Most perfectionists offer rival accounts of welfare rather than renounce welfare. But at least one – its leading contemporary exponent – claims to put no value on welfare.

Thomas Hurka is tempted by 'pure' perfectionism. He thinks pure perfectionism puts no value on welfare. Welfare or well-being is 'characterized subjectively, in terms of actual or hypothetical desires. Given this subjective characterization, perfectionism cannot concern well-being.'[39] So far, this allows perfectionists to value welfare when welfare is given a desire-independent construal. But Hurka resists any characterization of welfare 'not in terms of desire': 'I do not see that "developing human nature constitutes well-being and is therefore good" says anything over and above "developing human nature is good"'.[40] One can say that developing human nature is 'good for' the agent, but this says nothing more than that developing human nature is '(simply) good and a state of the person'.[41]

Hurka can be interpreted in two ways. The tough interpretation gives him an unattractive position. The soft interpretation concedes that even given pure perfectionism, welfare is valuable.

On the tough interpretation, Hurka makes no claims about perfection being good for the agent. The perfectionist speaks of '(simply) good', not 'good for', since 'good for' is infected with desire-dependence. Hurka

apart from its consequences.' Where desert or moral character enter, the presence of pleasure 'creates only a presumption that the total fact is good' (Ross, p. 137).

[39] Thomas Hurka, *Perfectionism* (Oxford, 1993), p. 17.

[40] Ibid., p. 194 n. 17.

[41] Ibid., p. 17. In '"Good" and "Good For"', *Mind* 96 (1987), pp. 72–3, Hurka argues that 'good for' should be 'banished' since it is ambiguous, but he again admits that it could mean 'that portion of the good . . . that falls within a person's own life'.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

should not welcome the tough interpretation, for perfectionists have often been motivated by reflection on what is good for the agent. When they argue (as Derek Parfit suggests) for the possibility that 'what someone would most want both now and later, fully knowing about the alternatives, would *not* be what would be best for him', they believe they are making a recommendation about what is best *for the agent* – not what is best, say, from an aesthetic or moral viewpoint.[42] They intend perfectionism to be a rival to preference satisfaction – but if preference satisfaction is a theory of what is good for one, and perfectionism a theory about what is (simply) good, there need be no conflict.

Perfectionists have a further reason for speaking of what is good for one. They need to fend off a worry about why we should care about perfection. Learning that perfection is good for one (or others) lets the critic see the attraction of perfection. Otherwise, perfection seems to have the same weak normative import as our aesthetic evaluations of one another. This is a prominent theme in the Aristotle literature: Aristotle's defenders see the need to show that performing one's function is good for us, since otherwise it is unclear what normative upshot functions play.

Hurka is willing to claim that perfection is good for the agent, provided this means 'perfection is good and a state of the agent'. On the soft interpretation, this analysis of 'good for' does not explain away the concept. It merely identifies the concept with another. One way to show that the concept is not explained away is to consider 'narrow' hedonists – hedonists who are not preference hedonists. They make a claim of the same form: 'pleasure is good and a state of the agent'. When they make this claim, no one thinks they have renounced talking of what is 'good for' the agent, although they have disconnected 'good for' and desires. On the soft interpretation, Hurka, like narrow hedonists, values what is good for agents. Nor does he disagree with the value of welfare when welfare is simply what is good for the agent.

Another way to show that 'good for' has not been explained away is to note that Hurka's perfection fits the truisms associated earlier with welfare. It has the same role as, say, hedonism, and so looks like a competing theory of welfare.

There is another argument for retaining the concept of 'good for', although this argument is not a defence of the soft interpretation. Hurka's identification of 'good for' with '(simply) good and a state of an agent' is implausible. Some things, such as narrow hedonist pleasures, are (simply) good, states of the agent, and good for the agent. But other things, such as a justified punishment, can be (simply) good and states

[42] Derek Parfit, *Reasons and Persons* (Oxford, 1984), p. 500.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

of the agent without being good for the agent.[43] We need an independent conception of what is 'good for' one. Perhaps nothing more can be said to specify this conception other than speaking of 'welfare' or 'well-being'. But this is not a problem: perfectionists who offer perfection as an account of welfare can defend the coherence of their project by noting that their claim is just like the claim of narrow hedonists – a claim we all understand.

I conclude that Hurka should prefer the soft to the tough interpretation. He should also prefer an independent sense of 'good for' to the identification offered on the soft interpretation. If so, he has no reason to resist the standard perfectionist view that perfection is a proposal about the nature of welfare. He should also count as valuing welfare.

5. The final objection holds that agreement on the goodness of welfare does not show that appeals to welfare require any less defence than appeals to other, more controversial, goods. Every proposition, whether contested or uncontested, requires defence.

I do not think this suggestion fits our epistemic practice. Perhaps any proposition might require defence; but at any one time, some require defence and some do not, and our agreement is what picks out those which do not. Disagreement is a guide to where epistemic energies should be directed. One might recommend a different epistemic practice, but then some rationale for this recommendation would be needed.

One might try to support a different practice by noting that uncontested claims can be wrong; defence is needed to preclude this threat. This possibility, however, can take two forms. Sometimes when we fear some claim is wrong, our fear is motivated – perhaps we have some evidence against the claim, or what appeared to be evidence for the claim now seems not to be good evidence, or the claim is inconsistent with a theory that does better. On the other hand, sometimes the fear that a claim is wrong is unmotivated – only the mere possibility of being wrong is noted. In the former case, we think defence is needed; the claim has become contested. In the latter, I do not think we see defence as fruitful. If so, the objection – that even uncontested claims need defence because we could be wrong – fails. For if there is a motivated worry about a claim, the claim is contested and so in need of defence; if there is no motivated worry, defence seems a waste of epistemic energies.[44]

[43] See Kagan, 'The Limits of Well-Being', *Social Philosophy and Policy* 9 (1992), p. 185.
[44] For the general approach, see J. L. Austin, 'Other Minds', in Austin, *Philosophical Papers* (Oxford, 1979).

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

III

Welfare needs no *defence*. Nonetheless, it would be reassuring to have an *explanation* of our agreement on its value, if only to dispel suspicion that the agreement is fragile or unreflective. Preference-based accounts of welfare can give such an explanation: We agree because part of what makes something a state of welfare is our positive evaluation of it. (Narrow hedonists and perfectionists need a different explanation.[45])

To illustrate, consider why appeals to welfare are so popular in anti-sceptical arguments. Consider, in particular, Thomas Nagel. In *The View from Nowhere*, Nagel asks what reasons we should admit. He starts with an uncontroversial example: I have an agent-relative reason to take aspirin because it will cure my headache.[46] But I also have an agent-neutral reason to take the aspirin (or give aspirin to you, should you have a headache). For the badness of the experience of the headache does not depend on some idiosyncratic feature of me. My reason depends simply on the feeling of pain – a feeling you too feel, when you have a headache. My reason, then, to take aspirin is also a reason for me to give you aspirin.[47]

Nagel implies an explanation for why pleasure and pain have this special status, and why they are the first and best choice for someone rejecting either scepticism about all justificatory reasons or scepticism about agent-neutral reasons. The goodness or badness of pleasures or pains is evident from the point of view of anyone capable of experience. This explains our agreement on their value.[48] And the goodness or badness of pleasures or pains is evident from any such point of view because our evaluation of the experience is partly constitutive of the experience as one of pleasure or pain. In the case of squeaking chalk, for example, the 'impersonal badness attaches not to the experience conceived merely as a certain sound, but to someone's having an experience he hates'.[49] (One might reply that constitution by our evaluation is irrelevant. Our evaluations might miss what is truly valuable. But (as Nagel suggests) it is hard to motivate this prospect,

[45] For one narrow hedonist explanation, see Peter Railton, 'Naturalism and Prescriptivity', *Social Philosophy and Policy* 7 (1989), p. 168.

[46] Parfit uses the same strategy of starting, against the sceptic, with an agent-relative reason for the agent to pursue his own welfare. 'Suppose that, unless I move, I shall be killed by a falling rock, and that what I now most want is to survive. Do I have a reason to move? It is undeniable that I do' (Parfit, *Reasons and Persons*, p. 452).

[47] Thomas Nagel, *The View from Nowhere* (Oxford, 1986), pp. 157–61.

[48] Hence Parfit's bold comment on the reason to move to avoid the rock: that I have a reason 'would not have been denied in any civilization at any time. This claim is *true*' (Parfit, *Reasons and Persons*, p. 452).

[49] Nagel, *The View from Nowhere*, p. 169, emphasis removed. This (preference-hedonist) view of pleasure and pain is common. See, for example, Sidgwick, *The Methods of Ethics*, pp. 127, 131; Parfit, *Reasons and Persons*, p. 501.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

at least when the evaluations are conducted in light of our best information.)

Of course Nagel rejects welfarism. For he supposes things other than pleasure or pain are good or bad from any point of view: freedom, self-respect, access to opportunities and resources.[50] It is significant, however, that Nagel does not try to establish the existence of agent-relative or agent-neutral reasons by appeal to these items. This is because, I think, none of these items is even partly constituted by our evaluation of it, and so it becomes at least possible to imagine a point of view from which the item lacks value. Agreement on the value of these goods is not secured in the way agreement on the value of pleasure or pain is secured. Thus, for example, it is possible for one to object, against the argument for agent-neutral reasons, that the badness of my lacking some resource (such as a high income) depends on some idiosyncratic feature of me (such as an expensive medical condition), and so my reason for getting the resource is not also a reason for me to give you the resource. This is what underwrites the familiar complaint, against primary goods or resources, that the effect of the resource on the recipient is wrongly omitted.[51]

The point can be put another way. Nagel is anxious to prevent desires *per se* from producing agent-neutral reasons. Your desire to climb Mount Kilimanjaro gives you an agent-relative reason to do so. It fails to provide an agent-neutral reason because the badness of failing to climb Kilimanjaro depends on an idiosyncratic feature of you – your desire to climb it. The badness of failing is not part of the failure; my failing is not bad at all, given that I have no desire to succeed. The badness of failing here seems on a par with the badness of lacking some resource or opportunity, in that one can lack the resource or opportunity without finding this bad. In this, pleasures and pains differ from idiosyncratic particular desires, resources and opportunities – and hence the attraction, to anyone arguing for agent-neutral (or agent-relative) reasons, of pleasures and pains. Of course, if there were true primary goods – all-purpose means – then such goods would be just as attractive in such arguments as pleasures and pains. But there seem to be no such means.

The emphasis on pleasure and pain might seem to restrict this explanation to preference hedonists about welfare. But pleasures and pains are not the only items partly constituted by our evaluations. For on Nagel's account, what makes pleasures and pains partly constituted by our evaluations is the role played, in picking them out, by our desires.

[50] Nagel, *The View from Nowhere*, p. 171.

[51] See, for example, Sen, 'Rights and Capabilities', in Honderich, *Morality and Objectivity*, p. 142.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press

And our desires can play the same role for things other than mental states. Your desire to climb Mount Kilimanjaro is idiosyncratic to you, but your desire to satisfy your considered desires, whatever they might be, is not. This desire is part of any (human) point of view, and the goodness of desire satisfaction is hence evident from any such point of view.[52] The explanation, then, covers preference satisfaction theorists as well as preference hedonists about welfare.

There is some distance from welfarism to utilitarianism. Arguments for agent-neutrality and maximizing aggregate welfare are needed. The argument for maximizing may be straightforward; anything other than maximizing seems to require admission of some other good. But my strategy does not look promising for agent-neutrality; disagreement is too widespread. Nor, for the same reason, does it look promising for deciding between rival theories of welfare. But in stressing the importance of welfarism to utilitarianism, I stress what the classic utilitarians themselves took to be most important. They took maximizing the aggregate virtually for granted. Only Sidgwick worries much about agent-neutrality. For all, the focus is on building a systematic moral theory around the value of welfare, rather than around more controversial 'intuitions' or 'caprice'. It is, I think, the appeal of this strategy that largely explains the appeal of utilitarianism.[53]

bshaver@cc.umanitoba.ca

[52] Nagel sees the point, but offers a bad objection. See Nagel, *The View from Nowhere*, p. 171 and, for a convincing reply, Sumner, *Welfare, Happiness, and Ethics*, pp. 197–8.

[53] I wish to thank Jack Bailey, Bob Bright, Roger Crisp, Elana Geller, Ken Gemes, Donovan Hulce, Iain Law, Cynthia Read, and, especially, Joyce Jenkins, along with audiences at the University of Manitoba Philosophy Department, University of Manitoba Philosophy Club, and at the 1998 Hume Conference in Stirling for comments or discussion.

https://doi.org/10.1017/S0953820804001153 Published online by Cambridge University Press