

Welfarism

Forthcoming in *International Encyclopedia of Ethics*
(ed. Hugh LaFollette)

Ben Bramble
b.e.bramble@gmail.com

ABSTRACT

Welfarism is a theory of value (or the good) *simpliciter*. Theories of value are fundamentally concerned with explaining what makes some possible worlds better than others. Welfarism is the view according to which the relative value of possible worlds is fully determined by how individuals are faring—or, in other words, by the facts about well-being that obtain—in these worlds. This entry begins by distinguishing between various forms of welfarism (pure vs. impure welfarism, and then narrow vs. wide welfarism). It then outlines some of the key attractions of welfarism. Finally, it surveys some of the most serious objections to welfarism (including malicious pleasure, the well-being of the wicked, great works of art, the non-identity problem, human extinction, transitivity of betterness, and distribution).

What is welfarism?

Welfarism is sometimes discussed as a theory of morality in general (Keller 2009). But it is most influential as a theory strictly of value (or the good) *simpliciter* (Sen 1979, Sumner 1996, and Holtug 1993). Theories of value are fundamentally concerned with explaining what makes some possible worlds (or complete histories of the universe from beginning to end) better—either in certain respects or all up—than others (see INTRINSIC VALUE; GOOD AND GOOD FOR). Such theories, on plausible assumptions, enable evaluations of things other than possible worlds—for example, actions, events, and public policies. Would it be good if X were to occur? Plausibly, this is determined just by whether the possible world resulting from X would be better in any respect than the possible world resulting from whatever alternative to X is under consideration.

Welfarism is the view according to which the relative value of possible worlds is fully determined by how individuals are faring—or, in other words, by the facts about well-being that obtain—in these worlds (see WELL-BEING).

There are two main kinds of welfarism, which differ based on how broadly “how individuals are faring” and “the facts about well-being” are understood. On *pure* welfarism, the facts in question concern just the *levels* of well-being of the individuals in the relevant worlds (i.e., *how* well or poorly off they are, or the *degrees* to which they are well or poorly off, in various ways). On *impure* welfarism, by contrast, these facts might also concern other things relating to their well-being, such as the *kinds* of well-

being in question (for example, what sorts of pleasures are being felt), the *worthiness* of these individuals to be well off, how well-being is *distributed* throughout the population, and so on. Pure welfarism implies that the only way things could be better is if some being or beings were in some ways better off. On impure welfarism, by contrast, there might be other ways things could be better—for example, by some being's better deserving the level of well-being she currently enjoys.

Welfarism comes also in a *narrow* and a *wide* variety. According to narrow welfarism, one possible world can be better or worse than another only if there are some individuals who are the same in *both*. If two possible worlds contain entirely different groups of individuals, or else none at all, then they are *incomparable* in value (see INCOMMENSURABILITY (AND INCOMPARABILITY)). By contrast, wide welfarism denies that identity is important in this way. According to it, for one world to be better than another it is sufficient that those in it are better off, *whatever* their identities.

It is useful to think of pure narrow welfarism as welfarism's paradigmatic form. On pure narrow welfarism, a respect in which one world is better than another is always just some respect in which an individual who exists in both worlds is better off in the former than in the latter.

It is also useful to think of welfarism as the thing that is added to consequentialism to arrive at utilitarianism (see CONSEQUENTIALISM; UTILITARIANISM). Consequentialism is the theory of morality according to which what makes an action morally required, wrong, or permissible, is just the contribution of this act (or, on some consequentialist theories, acts relevantly like it) to the value of worlds. Utilitarianism is consequentialism plus the view that the value of worlds is fully determined by the facts about well-being (in the above specified sense). When philosophers object to utilitarianism, it is sometimes just because they object to welfarism—some, that is, regard consequentialism as only (or most) plausible when it is paired with a non-welfarist theory of value (Railton 1984). Others hold the opposite view, that welfarism is what is plausible in utilitarianism (Shaver 2004).

In what follows, I will explain the appeal of welfarism, and survey some of the most serious objections to it.

The appeal of welfarism

There are two main attractions of welfarism. First, many people feel that without sentient beings nothing would matter—nothing, that is, could be good or bad, or go better or worse. Welfarism offers a simple and intuitive explanation of why this might be so. It might be so because without sentient beings, there could be no well-being, and value is (in the relevant sense) all about well-being.

Second, most of the things that seem most obviously good (for example, pleasure, good health, conflict resolution, economic prosperity, etc.) seem good chiefly because they make lives better for beings. Similarly, most of the things that seem most obviously bad (pain, injury, etc.) seem bad chiefly because they make lives worse for beings. In light of this, a natural hypothesis is that value is all about well-being (Sumner 1996). Welfarism has a unity that gives it great explanatory force.

Pure welfarism has additional attractions. Views other than it can, under certain circumstances, seem fetishistic. As Simon Keller says, “It should give a non-welfarist pause when she is presented with a case in which her view recommends that we make some individuals worse off, just so as to uphold a value that, in this case, is of no good to anyone” (Keller 2009, p. 91). Ask yourself how you would feel if some scarce resource were to be allocated in a way that benefited nobody, for the sake of some non-welfarist good, when it could easily have been used to benefit yourself or some loved one.

Note that impure welfarism, like non-welfarist views, is vulnerable to this charge of fetishism. As I’ve mentioned, on impure welfarism, things can go better in ways that do not involve anyone’s life going better for them. So, it allows for occasions on which what’s best is not best for anyone.

Objections to welfarism

Here, I will sketch some of the most serious objections to welfarism.

a. Malicious pleasure

Well-being comes in different *kinds*. That is, there are different ways in which people’s lives can go well or poorly for them. According to some philosophers, certain kinds of well-being are more valuable *simpliciter* than others. Some kinds might even fail to make the world go better at all. Some might even make it go worse. A famous example is *schadenfreude* or malicious pleasure—i.e., pleasure taken in the suffering or misfortune of others. Such pleasure, according to some people, is good for its subject in virtue of being pleasurable, but has no, or possibly even negative, value *simpliciter*.

Impure welfarists can accept that not all kinds of well-being are equally good *simpliciter* (see ORGANIC UNITIES). Malicious pleasure, they can say, has less, or even no value, *simpliciter*. But what can pure welfarists say? One influential response is to maintain that malicious pleasure is good for its subject, and so valuable *simpliciter*, but that we might feel otherwise because we are confusing a pleasure’s value *simpliciter* with its subject’s *moral character* (see MORAL CHARACTER; VIRTUE). Irwin Goldstein, for example, writes: “When people feel malicious pleasure, *they* are bad for feeling so... There are two values here: the bad in the moral offensiveness and the good in the pleasure’s self-benefiting character” (Goldstein 2003, p. 26).

A second option for a pure welfarist is to reject the assumption that malicious pleasure is good for one in the first place, or at least to hold that such pleasure is less good for one than other kinds. It might be, for example, that the intrinsic value for one of a pleasure depends in some way on the nature of its object. Fred Feldman, for example, suggests that pleasure taken in the ugly, false, or bestial, might be worth less for one than other kinds of pleasures (Feldman 2004).

b. The well-being of the wicked

Many of us feel that the value *simpliciter* of one’s well-being—*whatever* kind of well-being it is—depends on one’s moral character or degree of virtue. It is better that good

people do well than that bad people do. It might even be that when bad people do well, this has no, or perhaps only negative, value *simpliciter*. W.D. Ross, famously, writes: “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” (Ross 1930, p. 138).

Again, impure welfarism is specially designed to accommodate such worries. Impure welfarists can just accept that the value *simpliciter* of someone’s well-being depends on her moral character. But how can a pure welfarist respond? One option is to employ a similar strategy to Goldstein’s above—i.e., accept that the well-being of the wicked has as much value *simpliciter* as that of morally good people, and try to explain away our intuitions to the contrary. A different strategy is to argue that the wicked (for one reason or another) *cannot* flourish, and so when we feel that it is not good that they do well, it is not genuinely their doing well that we are reacting to (see EUDAIMONISM).

c. Great works of art

Some have claimed that the value of great works of art cannot be wholly accounted for in terms of their impact on human well-being (see AESTHETICS AND ETHICS). Susan Wolf, for example, argues that there is so much great art in existence that no individual work confers a net benefit on any individual (aside, perhaps, from on artists themselves): “If the excellent novel or film you actually enjoyed had not been available, or the painting you contemplated on your last trip to an art museum had been on loan to another gallery, there would have been another just as good and worthwhile that you could have read, watched, or pored over instead” (Wolf 2010, p. 48). But such works, Wolf says, are valuable nonetheless.

In response, a welfarist might reject Wolf’s assumption that something must confer a *net* benefit on somebody to count as valuable, and then argue that great works confer unique *kinds* of benefits on us (Bramble 2015). Without Van Gogh’s *Sunflowers*, many of us would be worse off *in one respect*, and this fact is sufficient, on welfarism properly conceived, to ground this work’s value. Moreover, it might be added, many of us find we have a special connection to *particular* great works. Given this, each great work might benefit some individuals *more* than the alternative works these individuals would have encountered had it not been created.

Wolf also questions whether welfarism can explain the contribution of great art to well-being—i.e., the huge value for us of coming into contact with *some great works at all*. If great works were valuable *simpliciter* only in virtue of providing human beings with pleasure or desire-fulfilment, she argues, then intuitively they would be far less good for us than they actually seem to be. Their value for us, in other words, seems dependent on their having *independent* or non-welfarist value.

d. The non-identity problem

Why is climate change bad? Welfarism is consistent with the most natural answer: because of the harm it will do to future beings. Oddly, however, welfarism has a hard

time explaining why it would be good to *prevent* climate change. Consider that any effective action on climate change would affect our daily lives in ways that would change, even if only slightly, the times at which we conceive our children. This in turn would change the sperm-egg combinations from which future generations will grow, and so change the *identities* of members of future generations. For this reason, it has been suggested, no effective action on climate change can benefit those threatened by climate change: those who will come into existence *if we do nothing*. The most such action could do is prevent these people from coming into existence, and substitute for them an entirely different group of individuals who will be better off. (See NON-IDENTITY; POPULATION; INTERGENERATIONAL ETHICS.)

Wide welfarists can respond that the identities of future people here are irrelevant (Parfit 1984). There is value in taking action on climate change because it would make things better for future people, *whoever* they will happen to be. But there is a problem for this response. In the climate change case, we are presuming that there will be roughly *the same number of people* in the two possible futures under consideration. But what if there will not be? What if we are choosing between two worlds with very different numbers of people? What if we are choosing, for example, between a world of relatively few people all of whom are very well off, and a world of somewhat more people, all of whom are somewhat less well off? In which world are “things better for people, whoever these people happen to be”? What is a wide welfarist to say?

A wide welfarist might say that things go best for people when *average* well-being is highest. But this seems to entail that for any world of people with very high well-being there is a better world of just one person with a slightly higher level of well-being. Alternatively, a wide welfarist might say that things go best for people when *total* well-being is highest. But this seems to entail that for any world of people with very high well-being there is a better world of vastly more people in which everyone has a life “barely worth living” (i.e., of very low, but still positive, well-being) (see REPUGNANT CONCLUSION; see PARFIT, DEREK).

What can a narrow welfarist say in response to the problem? She might say that even though effective action on climate change won’t make things better, failure to take it would still be morally wrong (for one reason or another) (see DEONTOLOGY; CLIMATE CHANGE). This moral wrongness might itself provide us with a reason to take action on climate change (see REASONS FOR ACTION, MORALITY AND).

e. Human extinction

Suppose a future generation of humans unanimously decides to be the final generation. None of these people desire to have children, or even to see humanity continue, even though any future people would certainly be well-off. Their decision seems very bad indeed. But who would it be bad for? It would not, it might be claimed, be bad for members of this final generation, since they would be getting what they wanted. It would also not be bad for future people, since there now will be no such people.

More starkly still, compare the actual world with one in which there never were any sentient beings. Surely the former is better, you might say. But how can welfarism account for this?

f. Transitivity of betterness

Many believe that the betterness relation is *transitive*—i.e., that if X is better than Y, and Y is better than Z, then Z must be better than X. But welfarism suggests it is not. Consider three possible worlds: A, B, and C. A contains only Adam (at welfare level 2) and Bill (at welfare level 1). B contains only Bill (at welfare level 2) and Cassie (at welfare level 1). C contains only Cassie (at welfare level 2) and Adam (at welfare level 1). Welfarism seems to imply that B is better than A, C is better than B, but that A is better than C.

Again, this is a problem chiefly for narrow welfarism. Wide welfarists can hold that A, B, and C are equally good. What can a narrow welfarist say? She might deny transitivity (Rachels 2001; Temkin 2012) (see INTRANSITIVITY).

g. Distribution

Suppose John is extremely poorly off, while Freddy is extremely well off. We can benefit only one of them. Who would it be better to benefit, other things equal? Many feel it would be better to benefit John. A welfarist can accept that this will be true in many cases, since in many cases, we will be able to benefit John *more* with some limited amount of resources. A hundred dollars, for example, will benefit people much more if it is spent on providing basic necessities to the world's poorest than if it is spent on a new pair of sneakers for a wealthy Westerner. But what if we *stipulate* that the proposed benefits to John and Freddy are proportionately great? A welfarist seems committed to saying that these options are equally good. This strikes many people as implausible.

Likewise, many feel that equality in living standards among citizens is a good-making feature of countries (see EQUALITY). Other things equal, it is better if citizens are roughly on a par with each other when it comes to their well-being than if there is substantial inequality among them. How can a welfarist explain this?

Again, impure welfarism has no difficulty accommodating such intuitions. Impure welfarists can simply grant that, say, equality in well-being is a good-making feature (see EGALITARIANISM; PRIORITARIANISM). But what can a pure welfarist say? She might cast doubt on the intuition that equality is intrinsically valuable by noting that increasing equality by reducing the well-being of the better off does not seem to make things better in *any* way. Suppose there are only two people: Freddy, who can see, and John, who is blind. It does not seem in any way good to blind Freddy.

See also: AESTHETICS AND ETHICS; CLIMATE CHANGE; CONSEQUENTIALISM; DEONTOLOGY; EGALITARIANISM; EQUALITY; EUDAIMONISM; INCOMMENSURABILITY (AND INCOMPARABILITY); INTERGENERATIONAL ETHICS; INTRANSITIVITY; INTRINSIC VALUE; GOOD AND GOOD FOR; MORAL CHARACTER; NON-IDENTITY; ORGANIC UNITIES; PARFIT, DEREK; POPULATION; PRIORITARIANISM; REASONS FOR ACTION, MORALITY AND; REPUGNANT CONCLUSION; UTILITARIANISM; VIRTUE; WELL-BEING.

REFERENCES

- Bramble, Ben 2015. “On Susan Wolf’s “Good-for-Nothings”” *Ethical Theory and Moral Practice*, vol. 18, pp. 1071-1081.
- Feldman, Fred 2004. *Pleasure and the Good Life*. New York: Oxford University Press.
- Goldstein, Irwin 2003. “Malicious pleasure evaluated: Is pleasure an unconditional good?” *Pacific Philosophical Quarterly*, vol. 84, pp. 24-31.
- Keller, Simon 2009. “Welfarism”. *Philosophy Compass* vol. 4, pp. 82–95.
- Holtug, Nils 2003. “Welfarism—The Very Idea”. *Utilitas* 15, pp. 151–74.
- Parfit, Derek 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Rachels, Stuart 2001. “A Set of Solutions to Parfit’s Problems.” *Nous*, vol. 35, pp. 214-238.
- Railton, Peter 1984. “Alienation, Consequentialism, and the Demands of Morality.” *Philosophy and Public Affairs*, vol. 13, pp. 134-171.
- Ross, William David 1930. *The Right and the Good*. Oxford University Press.
- Sen, Amartya 1979. “Utilitarianism and Welfarism”. *The Journal of Philosophy* vol. 76: pp. 463–89.
- Shaver, Robert 2004. “The Appeal of Utilitarianism”. *Utilitas* vol. 16, pp. 235–50.
- Sumner, Wayne 1996. *Welfare, Happiness and Ethics*. Oxford: Oxford UP.
- Temkin, Larry 2012. *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford University Press.
- Wolf, Susan 2010. “Good for Nothings” *Proceedings and Addresses of the American Philosophical Association* vol. 85, pp. 47-64.

FURTHER READINGS

- Adler, Matthew 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*, Oxford: Oxford University Press.
- Arneson, Richard 2002. “The End of Welfare As We Know It? Scanlon Versus Welfarist Consequentialism.” *Social Theory and Practice*. vol. 28, pp. 315-336.

Arrhenius, Gustaf 2009. "Can the Person Affecting Restriction Solve the Problems in Population Ethics?" in M. Roberts & D. Wasserman (eds.) *Harming Future Persons*, Ashgate, 2009.

Bentham, Jeremy 1789. *The Principles of Morals and Legislation*. New York: Hafner Publishing Co.

Bramble, Ben 2014. "Whole-Life Welfarism". *American Philosophical Quarterly*. vol. 51, pp. 63-74.

Crisp, Roger 2006. *Reasons and the Good*. Oxford University Press.

Fletcher, Guy 2008. "The Consistency of Qualitative Hedonism and the Value of (at least some) Malicious Pleasures". *Utilitas*, 20, pp. 462–471.

Glover, Jonathan 1977. *Causing Death and Saving Lives*, Penguin Books.

Kagan, Shelly 1997. *Normative Ethics*. Westview.

Moore, Andrew and Crisp, Roger 1996. "Welfarism in Moral Theory". *Australasian Journal of Philosophy*. vol. 74, pp. 598-613.

Moore, George Edward 1988. *Principia Ethica*. Amherst, NY: Prometheus Books.

Narveson, Jan. "Utilitarianism and New Generations". *Mind* 76, pp. 62-72.

Rabinowicz, Wlodek, and Rønnow-Rasmussen, Toni, "The Strike of the Demon: On Fitting Pro-attitudes and Value". *Ethics* 114, pp. 391-423.

Scanlon, Thomas 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press.

Sidgwick, Henry 1981. *The Methods of Ethics*. Indianapolis, IN: Hackett Publishing Co.