*Consequentialism*
**Walter Sinnott-Armstrong**
*(First published Tue May 20, 2003; substantive revision Thu Oct 22, 2015)*

Consequentialism, as its name suggests, is the view that normative properties depend only on consequences. This general approach can be applied at different levels to different normative properties of different kinds of things, but the most prominent example is consequentialism about the moral rightness of acts, which holds that whether an act is morally right depends only on the consequences of that act or of something related to that act, such as the motive behind the act or a general rule requiring acts of the same kind.

## 1. Classic Utilitarianism

The paradigm case of consequentialism is utilitarianism, whose classic proponents were Jeremy Bentham (1789), John Stuart Mill (1861), and Henry Sidgwick (1907). (For predecessors, see Schneewind 1990.) Classic utilitarians held hedonistic act consequentialism. Act consequentialism is the claim that an act is morally right if and only if that act maximizes the good, that is, if and only if the total amount of good for all minus the total amount of bad for all is greater than this net amount for any incompatible act available to the agent on that occasion. (Cf. Moore 1912, chs. 1–2.) Hedonism then claims that pleasure is the only intrinsic good and that pain is the only intrinsic bad.

These claims are often summarized in the slogan that an act is right if and only if it causes "the greatest happiness for the greatest number." This slogan is misleading, however. An act can increase happiness for most (the greatest number of) people but still fail to maximize the net good in the world if the smaller number of people whose happiness is not increased lose much more than the greater number gains. The principle of utility would not allow that kind of sacrifice of the smaller number to the greater number unless the net good overall is increased more than any alternative.

Classic utilitarianism is consequentialist as opposed to deontological because of what it denies. It denies that moral rightness depends directly on anything other than consequences, such as whether the agent promised in the past to do the act now. Of course, the fact that the agent promised to do the act might indirectly affect the act's consequences if breaking the promise will make other people unhappy. Nonetheless, according to classic utilitarianism, what makes it morally wrong to break the promise is its future effects on those other people rather than the fact that the agent promised in the past.

Since classic utilitarianism reduces all morally relevant factors (Kagan 1998, 17–22) to consequences, it might appear simple. However, classic utilitarianism is actually a complex combination of many distinct claims, including the following claims about the moral rightness of acts:

> Consequentialism = whether an act is morally right depends only on consequences (as opposed to the circumstances or the intrinsic nature of the act or anything that happens before the act).

> Actual Consequentialism = whether an act is morally right depends only on the actual consequences (as opposed to foreseen, foreseeable, intended, or likely consequences).

> Direct Consequentialism = whether an act is morally right depends only on the consequences of that act itself (as opposed to the consequences of the agent's motive, of a rule or practice that covers other acts of the same kind, and so on).

> Evaluative Consequentialism = moral rightness depends only on the value of the consequences (as opposed to non-evaluative features of the consequences).

> Hedonism = the value of the consequences depends only on the pleasures and pains in the consequences (as opposed to other supposed goods, such as freedom, knowledge, life, and so on).

> Maximizing Consequentialism = moral rightness depends only on which consequences are best (as

opposed to merely satisfactory or an improvement over the status quo).

Aggregative Consequentialism = which consequences are best is some function of the values of parts of those consequences (as opposed to rankings of whole worlds or sets of consequences).

Total Consequentialism = moral rightness depends only on the total net good in the consequences (as opposed to the average net good per person).

Universal Consequentialism = moral rightness depends on the consequences for all people or sentient beings (as opposed to only the individual agent, members of the individual's society, present people, or any other limited group).

Equal Consideration = in determining moral rightness, benefits to one person matter just as much as similar benefits to any other person (= all who count count equally).

Agent-neutrality = whether some consequences are better than others does not depend on whether the consequences are evaluated from the perspective of the agent (as opposed to an observer).

These claims could be clarified, supplemented, and subdivided further. What matters here is just that most pairs of these claims are logically independent, so a moral theorist could consistently accept some of them without accepting others. Yet classic utilitarians accepted them all. That fact makes classic utilitarianism a more complex theory than it might appear at first sight.

It also makes classic utilitarianism subject to attack from many angles. Persistent opponents posed plenty of problems for classic utilitarianism. Each objection led some utilitarians to give up some of the original claims of classic utilitarianism. By dropping one or more of those claims, descendants of utilitarianism can construct a wide variety of moral theories. Advocates of these theories often call them consequentialism rather than utilitarianism so that their theories will not be subject to refutation by association with the classic utilitarian theory.

## 2. What is Consequentialism?

This array of alternatives raises the question of which moral theories count as consequentialist (as opposed to deontological), and why. In actual usage, the term 'consequentialism' seems to be used as a family resemblance term to refer to any descendant of classic utilitarianism that remains close enough to its ancestor in the important respects. Of course, different philosophers see different respects as the important ones. Hence, there is no agreement on which theories count as consequentialist under this definition.

To resolve this vagueness, we need to determine which of the various claims of classic utilitarianism are essential to consequentialism. One claim seems clearly necessary. Any consequentialist theory must accept the claim that I labeled 'consequentialism', namely, that certain normative properties depend only on consequences. If that claim is dropped, the theory ceases to be consequentialist.

It is less clear whether that claim by itself is sufficient to make a theory consequentialist. Several philosophers assert that a moral theory should not be classified as consequentialist unless it is agent-neutral (McNaughton and Rawling 1991, Howard-Snyder 1994, Pettit 1997). This narrower definition is motivated by the fact that many self-styled critics of consequentialism argue against agent-neutrality.

Other philosophers prefer a broader definition that does not require a moral theory to be agent-neutral in order to be consequentialist (Bennett 1989; Broome 1991, 5–6; and Skorupski 1995). Criticisms of agent-neutrality can then be understood as directed against one part of classic utilitarianism that need not be adopted by every moral theory that is consequentialist. Moreover, according to those who prefer a broader definition of consequentialism, the narrower definition conflates independent claims and obscures a crucial commonality between agent-neutral consequentialism and other moral theories that focus exclusively on consequences, such as moral egoism and recent self-styled consequentialists who allow agent-relativity into their theories of value (Sen 1982, Broome 1991, Portmore 2001, 2003).

A definition solely in terms of consequences might seem too broad, because it includes absurd theories such as the theory that an act is morally right if it increases the number of goats in Texas. Of course, such theories are implausible. Still, it is not implausible to call them consequentialist, since they do look only at consequences. The implausibility of one version of consequentialism does not make consequentialism implausible in general, since other versions of consequentialism still might be plausible.

Besides, anyone who wants to pick out a smaller set of moral theories that excludes this absurd theory may talk about evaluative consequentialism, which is the claim that moral rightness depends only on the value of the consequences. Then those who want to talk about the even smaller group of moral theories that accepts both evaluative consequentialism and agent-neutrality may describe them as agent-neutral evaluative consequentialism. If anyone still insists on calling these smaller groups of theories by the simple name, 'consequentialism', this narrower usage will not affect any substantive issue.

Still, if the definition of consequentialism becomes too broad, it might seem to lose force. Some philosophers have argued that any moral theory, or at least any plausible moral theory, could be represented as a version of consequentialism (Sosa 1993, Portmore 2009, Dreier 1993 and 2011; but see Brown 2011). If so, then it means little to label a theory as consequentialist. The real content comes only by contrasting theories that are not consequentialist.

In the end, what matters is only that we get clear about which theories a particular commentator counts as consequentialist or not and which claims are supposed to make them consequentialist or not. Only then can we know which claims are at stake when this commentator supports or criticizes what they call "consequentialism". Then we can ask whether each objection really refutes that particular claim.

### 3. What is Good? Hedonistic vs. Pluralistic Consequentialisms
Some moral theorists seek a single simple basic principle because they assume that simplicity is needed in order to decide what is right when less basic principles or reasons conflict. This assumption seems to make hedonism attractive. Unfortunately, however, hedonism is not as simple as they assume, because hedonists count both pleasures and pains. Pleasure is distinct from the absence of pain, and pain is distinct from the absence of pleasure, since people sometimes feel neither pleasure nor pain, and sometimes they feel both at once. Nonetheless, hedonism was adopted partly because it seemed simpler than competing views.

The simplicity of hedonism was also a source of opposition. From the start, the hedonism in classic utilitarianism was treated with contempt. Some contemporaries of Bentham and Mill argued that hedonism lowers the value of human life to the level of animals, because it implies that, as Bentham said, an unsophisticated game (such as push-pin) is as good as highly intellectual poetry if the game creates as much pleasure (Bentham 1843). Quantitative hedonists sometimes respond that great poetry almost always creates more pleasure than trivial games (or sex and drugs and rock-and-roll), because the pleasures of poetry are more certain, durable, fecund, and so on.

Mill used a different strategy to avoid calling push-pin as good as poetry. He distinguished higher and lower qualities of pleasures according to the preferences of people who have experienced both kinds (Mill 1861, 56; compare Plato 1993 and Hutcheson 1755, 421–23). This qualitative hedonism has been subjected to much criticism, including charges that it is incoherent and does not count as hedonism (Moore 1903, 80–81; cf. Feldman 1997, 106–24).

Even if qualitative hedonism is coherent and is a kind of hedonism, it still might not seem plausible. Some critics argue that not all pleasures are valuable, since, for example, there is no value in the pleasures that a sadist gets from whipping a victim or that an addict gets from drugs. Other opponents object that not only pleasures are intrinsically valuable, because other things are valuable independently of whether they lead to pleasure or avoid pain. For example, my love for my wife does not seem to become less valuable when I get less pleasure from her because she contracts some horrible disease. Similarly, freedom seems valuable even when it creates anxiety, and even when it is freedom to do something (such as leave one's country) that one does not want to do. Again, many people value knowledge of other galaxies regardless of whether this knowledge will create pleasure or avoid pain.

These points against hedonism are often supplemented with the story of the experience machine found in Nozick (1974, 42–45; cf. the movie, The Matrix). People on this machine believe they are spending time with their friends, winning Olympic gold medals and Nobel prizes, having sex with their favorite lovers, or doing whatever gives them the greatest balance of pleasure over pain. Although they have no real friends or lovers and actually accomplish nothing, people on the experience machine get just as much pleasure as if their beliefs were true. Moreover, they feel no (or little) pain. Assuming that the machine is reliable, it would seem irrational not to hook oneself up to this machine if pleasure and pain were all that mattered, as hedonists claim. Since it does not seem irrational to refuse to hook oneself up to this machine, hedonism seems inadequate. The reason is that hedonism overlooks the value of real friendship, knowledge, freedom, and achievements, all of which are lacking for deluded people on the experience machine.

Some hedonists claim that this objection rests on a misinterpretation of hedonism. If hedonists see pleasure and pain as sensations, then a machine might be able to reproduce those sensations. However, we can also say that a mother is pleased that her daughter gets good grades. Such propositional pleasure occurs only when the state of affairs in which the person takes pleasure exists (that is, when the daughter actually gets good grades). But the relevant states of affairs would not really exist if one were hooked up to the experience machine. Hence, hedonists who value propositional pleasure rather than sensational pleasure can deny that more pleasure is achieved by hooking oneself up to such an experience machine (Feldman 1997, 79–105; see also Tannsjo 1998 and Feldman 2004 for more on hedonism).

A related position rests on the claim that what is good is desire satisfaction or the fulfillment of preferences; and what is bad is the frustration of desires or preferences. What is desired or preferred is usually not a sensation but is, rather, a state of affairs, such as having a friend or accomplishing a goal. If a person desires or prefers to have true friends and true accomplishments and not to be deluded, then hooking this person up to the experience machine need not maximize desire satisfaction. Utilitarians who adopt this theory of value can then claim that an agent morally ought to do an act if and only if that act maximizes desire satisfaction or preference fulfillment, regardless of whether the act causes sensations of pleasure. This position is usually described as preference utilitarianism.

One problem for preference utilitarianism concerns how to make interpersonal comparisons (though this problem also arises for some other theories of value). If we want to know what one person prefers, we can ask what that person would choose in conflicts. We cannot, however, use the same method to determine whether one person's preference is stronger or weaker than another person's preference, since these different people might choose differently in the decisive conflicts. We need to settle which preference (or pleasure) is stronger because we may know that Jones prefers A's being done to A's not being done (and Jones would receive more pleasure from A's being done than from A's not being done), whereas Smith prefers A's not being done (and Smith would receive more pleasure from A's not being done than from A's being done). To determine whether it is right to do A or not to do A, we must be able to compare the strengths of Jones's and Smith's preferences (or the amounts of pleasure each would receive in her preferred outcome) in order to determine whether doing A or not doing A would be better overall. Utilitarians and consequentialists have proposed many ways to solve this problem of interpersonal comparison, and each attempt has received criticisms. Debates about this problem still rage. (For a recent discussion with references, see Coakley 2015.)

Preference utilitarianism is also often criticized on the grounds that some preferences are misinformed, crazy, horrendous, or trivial. I might prefer to drink the liquid in a glass because I think that it is beer, though it really is strong acid. Or I might prefer to die merely because I am clinically depressed. Or I might prefer to torture children. Or I might prefer to spend my life learning to write as small as possible. In all such cases, opponents of preference utilitarianism can deny that what I prefer is really good. Preference utilitarians can respond by limiting the preferences that make something good, such as by referring to informed desires that do not disappear after therapy (Brandt 1979). However, it is not clear that such qualifications can solve all of the problems for a preference theory of value without making the theory circular by depending on substantive assumptions about which preferences are for good things.

Many consequentialists deny that all values can be reduced to any single ground, such as pleasure or desire satisfaction, so they instead adopt a pluralistic theory of value. Moore's ideal utilitarianism, for example, takes into account the values of beauty and truth (or knowledge) in addition to pleasure (Moore 1903, 83–85, 194;

1912). Other consequentialists add the intrinsic values of friendship or love, freedom or ability, life, virtue, and so on.

If the recognized values all concern individual welfare, then the theory of value can be called welfarist (Sen 1979). When a welfarist theory of value is combined with the other elements of classic utilitarianism, the resulting theory can be called welfarist consequentialism.

One non-welfarist theory of value is perfectionism, which claims that certain states make a person's life good without necessarily being good for the person in any way that increases that person's welfare (Hurka 1993, esp. 17). If this theory of value is combined with other elements of classic utilitarianism, the resulting theory can be called perfectionist consequentialism or, in deference to its Aristotelian roots, eudaemonistic consequentialism.

Similarly, some consequentialists hold that an act is right if and only if it maximizes some function of both happiness and capabilities (Sen 1985, Nussbaum 2000). Disabilities are then seen as bad regardless of whether they are accompanied by pain or loss of pleasure.

Or one could hold that an act is right if it maximizes respect for (or minimizes violations of) certain specified moral rights. Such theories are sometimes described as a utilitarianism of rights. This approach could be built into total consequentialism with rights weighed against happiness and other values or, alternatively, the disvalue of rights violations could be lexically ranked prior to any other kind of loss or harm (cf. Rawls 1971, 42). Such a lexical ranking within a consequentialist moral theory would yield the result that nobody is ever justified in violating rights for the sake of happiness or any value other than rights, although it would still allow some rights violations in order to avoid or prevent other rights violations.

When consequentialists incorporate a variety of values, they need to rank or weigh each value against the others. This is often difficult. Some consequentialists even hold that certain values are incommensurable or incomparable in that no comparison of their values is possible (Griffin 1986 and Chang 1997). This position allows consequentialists to recognize the possibility of irresolvable moral dilemmas (Sinnott-Armstrong 1988, 81; Railton 2003, 249-91).

Pluralism about values also enables consequentialists to handle many of the problems that plague hedonistic utilitarianism. For example, opponents often charge that classical utilitarians cannot explain our obligations to keep promises and not to lie when no pain is caused or pleasure is lost. Whether or not hedonists can meet this challenge, pluralists can hold that knowledge is intrinsically good and/or that false belief is intrinsically bad. Then, if deception causes false beliefs, deception is instrumentally bad, and agents ought not to lie without a good reason, even when lying causes no pain or loss of pleasure. Since lying is an attempt to deceive, to lie is to attempt to do what is morally wrong (in the absence of defeating factors). Similarly, if a promise to do an act is an attempt to make an audience believe that the promiser will do the act, then to break a promise is for a promiser to make false a belief that the promiser created. Although there is more tale to tell, the disvalue of false belief can be part of a consequentialist story about why it is morally wrong to break promises.

When such pluralist versions of consequentialism are not welfarist, some philosophers would not call them utilitarian. However, this usage is not uniform, since even non-welfarist views are sometimes called utilitarian. Whatever you call them, the important point is that consequentialism and the other elements of classical utilitarianism are compatible with many different theories about which things are good or valuable.

Instead of turning pluralist, some consequentialists foreswear the aggregation of values. Classic utilitarianism added up the values within each part of the consequences to determine which total set of consequences has the most value in it. One could, instead, aggregate goods for each individual but not aggregate goods of separate individuals (Roberts 2002). Or one could give up aggregation altogether and just rank total sets of consequences or total worlds created by acts without breaking those worlds down into valuable parts. One motive for this move is Moore's principle of organic unity (Moore 1903, 27–36). For example, even if punishment of a criminal causes pain, a consequentialist can hold that a world with both the crime and the punishment is better than a world with the crime but not the punishment, perhaps because the former contains more justice. Similarly, a world might seem better when people do not get pleasures that they do not deserve. Cases like these lead some consequentialists to deny that moral rightness is any function of the values of particular effects of acts. Instead,

they compare the whole world (or total set of consequences) that results from an action with the whole world that results from not doing that action. If the former is better, then the action is morally right (J.J.C. Smart 1973, 32; Feldman 1997, 17–35). This approach can be called holistic consequentialism or world utilitarianism.

Another way to incorporate relations among values is to consider distribution. Compare one outcome where most people are destitute but a few lucky people have extremely large amounts of goods with another outcome that contains slightly less total goods but where every person has nearly the same amount of goods. Egalitarian critics of classical utilitarianism argue that the latter outcome is better, so more than the total amount of good matters. Traditional hedonistic utilitarians who prefer the latter outcome often try to justify egalitarian distributions of goods by appealing to a principle of diminishing marginal utility. Other consequentialists, however, incorporate a more robust commitment to equality. Early on, Sidgwick (1907, 417) responded to such objections by allowing distribution to break ties between other values. More recently, some consequentialists have added some notion of fairness (Broome 1991, 192–200) or desert (Feldman 1997, 154–74) to their test of which outcome is best. (See also Kagan 1998, 48–59.) Such consequentialists do not just add up values; they look at patterns.

A related issue arises from population change. Imagine that a government considers whether to provide free contraceptives to curb a rise in population. Without free contraceptives, overcrowding will bring hunger, disease, and pain, so each person will be worse off. Still, each new person will have enough pleasure and other goods that the total net utility will increase with the population. Classic utilitarianism focuses on total utility, so it seems to imply that this government should not provide free contraceptives. That seems implausible to many utilitarians. To avoid this result, some utilitarians claim that an act is morally wrong if and only if its consequences contain more pain (or other disvalues) than an alternative, regardless of positive values. This negative utilitarianism implies that the government should provide contraceptives, since that program reduces pain (and other disvalues), even though it also decreases total net pleasure (or good). Unfortunately, negative utilitarianism also seems to imply that the government should painlessly kill everyone it can, since dead people feel no pain (and have no false beliefs, diseases, or disabilities – though killing them does cause loss of ability) (cf. R.N. Smart 1958). A more popular response is average utilitarianism, which says that the best consequences are those with the highest average utility (cf. Rawls 1971, 161–75). The average utility would be higher with the contraceptive program than without it, so average utilitarianism yields the more plausible result—that the government should adopt the contraceptive program. Critics sometimes charge that the average utility could also be increased by killing the worst off, but this claim is not at all clear, because such killing would put everyone in danger (since, after the worst off are killed, another group becomes the worst off, and then they might be killed next). Still, average utilitarianism faces problems of its own (such as "the mere addition paradox" in Parfit 1984, chap. 19). In any case, all maximizing consequentialists, whether or not they are pluralists, must decide whether moral rightness depends on maximizing total good or average good.

A final challenge to consequentialists' accounts of value derives from Geach 1956 and has been pressed recently by Thomson 2001. Thomson argues that "A is a good X" (such as a good poison) does not entail "A is good", so the term "good" is an attributive adjective and cannot legitimately be used without qualification. On this view, it is senseless to call something good unless this means that it is good for someone or in some respect or for some use or at some activity or as an instance of some kind. Consequentialists are supposed to violate this restriction when they say that the total or average consequences or the world as a whole is good without any such qualification. However, consequentialists can respond either that the term "good" has predicative uses in addition to its attributive uses or that when they call a world or total set of consequences good, they are calling it good for consequences or for a world (Sinnott-Armstrong 2003a). If so, the fact that "good" is often used attributively creates no problem for consequentialists.

## 4. Which Consequences? Actual vs. Expected Consequentialisms

A second set of problems for classic utilitarianism is epistemological. Classic utilitarianism seems to require that agents calculate all consequences of each act for every person for all time. That's impossible.

This objection rests on a misinterpretation. Critics assume that the principle of utility is supposed to be used as a decision procedure or guide, that is, as a method that agents consciously apply to acts in advance to help them make decisions. However, most classic and contemporary utilitarians and consequentialists do not propose their principles as decision procedures. (Bales 1971) Bentham wrote, "It is not to be expected that this process [his

hedonic calculus] should be strictly pursued previously to every moral judgment." (1789, Chap. IV, Sec. VI) Mill agreed, "it is a misapprehension of the utilitarian mode of thought to conceive it as implying that people should fix their minds upon so wide a generality as the world, or society at large." (1861, Chap. II, Par. 19) Sidgwick added, "It is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim." (1907, 413)

Instead, most consequentialists claim that overall utility is the criterion or standard of what is morally right or morally ought to be done. Their theories are intended to spell out the necessary and sufficient conditions for an act to be morally right, regardless of whether the agent can tell in advance whether those conditions are met. Just as the laws of physics govern golf ball flight, but golfers need not calculate physical forces while planning shots; so overall utility can determine which decisions are morally right, even if agents need not calculate utilities while making decisions. If the principle of utility is used as a criterion of the right rather than as a decision procedure, then classical utilitarianism does not require that anyone know the total consequences of anything before making a decision.

Furthermore, a utilitarian criterion of right implies that it would not be morally right to use the principle of utility as a decision procedure in cases where it would not maximize utility to try to calculate utilities before acting. Utilitarians regularly argue that most people in most circumstances ought not to try to calculate utilities, because they are too likely to make serious miscalculations that will lead them to perform actions that reduce utility. It is even possible to hold that most agents usually ought to follow their moral intuitions, because these intuitions evolved to lead us to perform acts that maximize utility, at least in likely circumstances (Hare 1981, 46–47). Some utilitarians (Sidgwick 1907, 489–90) suggest that a utilitarian decision procedure may be adopted as an esoteric morality by an elite group that is better at calculating utilities, but utilitarians can, instead, hold that nobody should use the principle of utility as a decision procedure.

This move is supposed to make consequentialism self-refuting, according to some opponents. However, there is nothing incoherent about proposing a decision procedure that is separate from one's criterion of the right. Similar distinctions apply in other normative realms. The criterion of a good stock investment is its total return, but the best decision procedure still might be to reduce risk by buying an index fund or blue-chip stocks. Criteria can, thus, be self-effacing without being self-refuting (Parfit 1984, chs. 1 and 4).

Others object that this move takes the force out of consequentialism, because it leads agents to ignore consequentialism when they make real decisions. However, a criterion of the right can be useful at a higher level by helping us choose among available decision procedures and refine our decision procedures as circumstances change and we gain more experience and knowledge. Hence, most consequentialists do not mind giving up consequentialism as a direct decision procedure as long as consequences remain the criterion of rightness (but see Chappell 2001).

If overall utility is the criterion of moral rightness, then it might seem that nobody could know what is morally right. If so, classical utilitarianism leads to moral skepticism. However, utilitarians insist that we can have strong reasons to believe that certain acts reduce utility, even if we have not yet inspected or predicted every consequence of those acts. For example, in normal circumstances, if someone were to torture and kill his children, it is possible that this would maximize utility, but that is very unlikely. Maybe they would have grown up to be mass murders, but it is at least as likely that they would cure serious diseases or do other great things, and it is much more likely that they would have led normally happy (or at least not destructive) lives. So observers as well as agents have adequate reasons to believe that such acts are morally wrong, according to act utilitarianism. In many other cases, it will still be hard to tell whether an act will maximize utility, but that shows only that there are severe limits to our knowledge of what is morally right. That should be neither surprising nor problematic for utilitarians.

If utilitarians want their theory to allow more moral knowledge, they can make a different kind of move by turning from actual consequences to expected or expectable consequences. Suppose that Alice finds a runaway teenager who asks for money to get home. Alice wants to help and reasonably believes that buying a bus ticket home for this runaway will help, so she buys a bus ticket and puts the runaway on the bus. Unfortunately, the bus is involved in a freak accident, and the runaway is killed. If actual consequences are what determine moral wrongness, then it was morally wrong for Alice to buy the bus ticket for this runaway. Opponents claim that this

result is absurd enough to refute classic utilitarianism.

Some utilitarians bite the bullet and say that Alice's act was morally wrong, but it was blameless wrongdoing, because her motives were good, and she was not responsible, given that she could not have foreseen that her act would cause harm. Since this theory makes actual consequences determine moral rightness, it can be called actual consequentialism.

Other responses claim that moral rightness depends on foreseen, foreseeable, intended, or likely consequences, rather than actual ones. Imagine that Bob does not in fact foresee a bad consequence that would make his act wrong if he did foresee it, but that Bob could easily have foreseen this bad consequence if he had been paying attention. Maybe he does not notice the rot on the hamburger he feeds to his kids which makes them sick. If foreseen consequences are what matter, then Bob's act is not morally wrong. If foreseeable consequences are what matter, then Bob's act is morally wrong, because the bad consequences were foreseeable. Now consider Bob's wife, Carol, who notices that the meat is rotten but does not want to have to buy more, so she feeds it to her children anyway, hoping that it will not make them sick; but it does. Carol's act is morally wrong if foreseen or foreseeable consequences are what matter, but not if what matter are intended consequences, because she does not intend to make her children sick. Finally, consider Bob and Carol's son Don, who does not know enough about food to be able to know that eating rotten meat can make people sick. If Don feeds the rotten meat to his little sister, and it makes her sick, then the bad consequences are not intended, foreseen, or even foreseeable by Don, but those bad results are still objectively likely or probable, unlike the case of Alice. Some philosophers deny that probability can be fully objective, but at least the consequences here are foreseeable by others who are more informed than Don can be at the time. For Don to feed the rotten meat to his sister is, therefore, morally wrong if likely consequences are what matter, but not morally wrong if what matter are foreseen or foreseeable or intended consequences.

Consequentialist moral theories that focus on actual or objectively probable consequences are often described as objective consequentialism (Railton 1984). In contrast, consequentialist moral theories that focus on intended or foreseen consequences are usually described as subjective consequentialism. Consequentialist moral theories that focus on reasonably foreseeable consequences are then not subjective insofar as they do not depend on anything inside the actual subject's mind, but they are subjective insofar as they do depend on which consequences this particular subject would foresee if he or she were better informed or more rational.

One final solution to these epistemological problems deploys the legal notion of proximate cause. If consequentialists define consequences in terms of what is caused (unlike Sosa 1993), then which future events count as consequences is affected by which notion of causation is used to define consequences. Suppose I give a set of steak knives to a friend. Unforeseeably, when she opens my present, the decorative pattern on the knives somehow reminds her of something horrible that her husband did. This memory makes her so angry that she voluntarily stabs and kills him with one of the knives. She would not have killed her husband if I had given her spoons instead of knives. Did my decision or my act of giving her knives cause her husband's death? Most people (and the law) would say that the cause was her act, not mine. Why? One explanation is that her voluntary act intervened in the causal chain between my act and her husband's death. Moreover, even if she did not voluntarily kill him, but instead she slipped and fell on the knives, thereby killing herself, my gift would still not be a cause of her death, because the coincidence of her falling intervened between my act and her death. The point is that, when voluntary acts and coincidences intervene in certain causal chains, then the results are not seen as caused by the acts further back in the chain of necessary conditions (Hart and Honoré 1985). Now, if we assume that an act must be such a proximate cause of a harm in order for that harm to be a consequence of that act, then consequentialists can claim that the moral rightness of that act is determined only by such proximate consequences. This position, which might be called proximate consequentialism, makes it much easier for agents and observers to justify moral judgments of acts because it obviates the need to predict non-proximate consequences in distant times and places. Hence, this move is worth considering, even though it has never been developed and deviates far from traditional consequentialism, which counts not only proximate consequences but all upshots — that is, everything for which the act is a causally necessary condition.

## 5. Consequences of What? Rights, Relativity, and Rules
Another problem for utilitarianism is that it seems to overlook justice and rights. One common illustration is called Transplant. Imagine that each of five patients in a hospital will die without an organ transplant. The patient

in Room 1 needs a heart, the patient in Room 2 needs a liver, the patient in Room 3 needs a kidney, and so on. The person in Room 6 is in the hospital for routine tests. Luckily (for them, not for him!), his tissue is compatible with the other five patients, and a specialist is available to transplant his organs into the other five. This operation would save their lives, while killing the "donor". There is no other way to save any of the other five patients (Foot 1966, Thomson 1976; compare related cases in Carritt 1947 and McCloskey 1965).

We need to add that the organ recipients will emerge healthy, the source of the organs will remain secret, the doctor won't be caught or punished for cutting up the "donor", and the doctor knows all of this to a high degree of probability (despite the fact that many others will help in the operation). Still, with the right details filled in, it looks as if cutting up the "donor" will maximize utility, since five lives have more utility than one life (assuming that the five lives do not contribute too much to overpopulation). If so, then classical utilitarianism implies that it would not be morally wrong for the doctor to perform the transplant and even that it would be morally wrong for the doctor not to perform the transplant. Most people find this result abominable. They take this example to show how bad it can be when utilitarians overlook individual rights, such as the unwilling donor's right to life.

Utilitarians can bite the bullet, again. They can deny that it is morally wrong to cut up the "donor" in these circumstances. Of course, doctors still should not cut up their patients in anything close to normal circumstances, but this example is so abnormal that we should not expect our normal moral rules to apply, and we should not trust our moral intuitions, which evolved to fit normal situations (Sprigge 1965). Many utilitarians are happy to reject common moral intuitions in this case, like many others (cf. Singer 1974, Unger 1996, Norcross 1997).

Most utilitarians lack such strong stomachs (or teeth), so they modify utilitarianism to bring it in line with common moral intuitions, including the intuition that doctors should not cut up innocent patients. One attempt claims that a killing is worse than a death. The doctor would have to kill the "donor" in order to prevent the deaths of the five patients, but nobody is killed if the five patients die. If one killing is worse than five deaths that do not involve killing, then the world that results from the doctor performing the transplant is worse than the world that results from the doctor not performing the transplant. With this new theory of value, consequentialists can agree with others that it is morally wrong for the doctor to cut up the "donor" in this example.

A modified example still seems problematic. Just suppose that the five patients need a kidney, a lung, a heart, and so forth because they were all victims of murder attempts. Then the world will contain the five killings of them if they die, but not if they do not die. Thus, even if killings are worse than deaths that are not killings, the world will still be better overall (because it will contain fewer killings as well as fewer deaths) if the doctor cuts up the "donor" to save the five other patients. But most people still think it would be morally wrong for the doctor to kill the one to prevent the five killings. The reason is that it is not the doctor who kills the five, and the doctor's duty seems to be to reduce the amount of killing that she herself does. In this view, the doctor is not required to promote life or decrease death or even decrease killing by other people. The doctor is, instead, required to honor the value of life by not causing loss of life (cf. Pettit 1997).

This kind of case leads some consequentialists to introduce agent-relativity into their theory of value (Sen 1982, Broome 1991, Portmore 2001, 2003). To apply a consequentialist moral theory, we need to compare the world with the transplant to the world without the transplant. If this comparative evaluation must be agent-neutral, then, if an observer judges that the world with the transplant is better, the agent must make the same judgment, or else one of them is mistaken. However, if such evaluations can be agent-relative, then it could be legitimate for an observer to judge that the world with the transplant is better (since it contains fewer killings by anyone), while it is also legitimate for the doctor as agent to judge that the world with the transplant is worse (because it includes a killing by him). In other cases, such as competitions, it might maximize the good from an agent's perspective to do an act, while maximizing the good from an observer's perspective to stop the agent from doing that very act. If such agent-relative value makes sense, then it can be built into consequentialism to produce the claim that an act is morally wrong if and only if the act's consequences include less overall value from the perspective of the agent. This agent-relative consequentialism, plus the claim that the world with the transplant is worse from the perspective of the doctor, could justify the doctor's judgment that it would be morally wrong for him to perform the transplant. A key move here is to adopt the agent's perspective in judging the agent's act. Agent-neutral consequentialists judge all acts from the observer's perspective, so they would judge the doctor's act to be wrong, since the world with the transplant is better from an observer's perspective. In contrast, an

agent-relative approach requires observers to adopt the doctor's perspective in judging whether it would be morally wrong for the doctor to perform the transplant. This kind of agent-relative consequentialism is then supposed to capture commonsense moral intuitions in such cases.

Agent-relativity is also supposed to solve other problems. W. D. Ross (1930, 34–35) argued that, if breaking a promise created only slightly more happiness overall than keeping the promise, then the agent morally ought to break the promise according to classic utilitarianism. This supposed counterexample cannot be avoided simply by claiming that keeping promises has agent-neutral value, since keeping one promise might prevent someone else from keeping another promise. Still, agent-relative consequentialists can respond that keeping a promise has great value from the perspective of the agent who made the promise and chooses whether or not to keep it, so the world where a promise is kept is better from the agent's perspective than another world where the promise is not kept, unless enough other values override the value of keeping the promise. In this way, agent-relative consequentialists can explain why agents morally ought not to break their promises in just the kind of case that Ross raised.

Similarly, critics of utilitarianism often argue that utilitarians cannot be good friends, because a good friend places more weight on the welfare of his or her friends than on the welfare of strangers, but utilitarianism requires impartiality among all people. However, agent-relative consequentialists can assign more weight to the welfare of a friend of an agent when assessing the value of the consequences of that agent's acts. In this way, consequentialists try to capture common moral intuitions about the duties of friendship (see also Jackson 1991).

One final variation still causes trouble. Imagine that the doctor herself wounded the five people who need organs. If the doctor does not save their lives, then she will have killed them herself. In this case, even if the doctor can disvalue killings by herself more than killings by other people, the world still seems better from her own perspective if she performs the transplant. Critics will object that it is, nonetheless, morally wrong for the doctor to perform the transplant. Many people will not find this intuition as clear as in the other cases, but those who do find it immoral for the doctor to perform the transplant even in this case will want to modify consequentialism in some other way in order to yield the desired judgment.

This problem cannot be solved by building rights or fairness or desert into the theory of value. The five do not deserve to die, and they do deserve their lives, just as much as the one does. Each option violates someone's right not to be killed and is unfair to someone. So consequentialists need more than just new values if they want to avoid endorsing this transplant.

One option is to go indirect. A direct consequentialist holds that the moral qualities of something depend only on the consequences of that very thing. Thus, a direct consequentialist about motives holds that the moral qualities of a motive depend on the consequences of that motive. A direct consequentialist about virtues holds that the moral qualities of a character trait (such as whether or not it is a moral virtue) depend on the consequences of that trait (Driver 2001a, Hurka 2001, Jamieson 2005, Bradley 2005). A direct consequentialist about acts holds that the moral qualities of an act depend on the consequences of that act. Someone who adopts direct consequentialism about everything is a global direct consequentialist (Pettit and Smith 2000).

In contrast, an indirect consequentialist holds that the moral qualities of something depend on the consequences of something else. One indirect version of consequentialism is motive consequentialism, which claims that the moral qualities of an act depend on the consequences of the motive of that act (compare Adams 1976 and Sverdlik 2011). Another indirect version is virtue consequentialism, which holds that whether an act is morally right depends on whether it stems from or expresses a state of character that maximizes good consequences and, hence, is a virtue.

The most common indirect consequentialism is rule consequentialism, which makes the moral rightness of an act depend on the consequences of a rule. Since a rule is an abstract entity, a rule by itself strictly has no consequences. Still, obedience rule consequentialists can ask what would happen if everybody obeyed a rule or what would happen if everybody violated a rule. They might argue, for example, that theft is morally wrong because it would be disastrous if everybody broke a rule against theft. Often, however, it does not seem morally wrong to break a rule even though it would cause disaster if everybody broke it. For example, if everybody broke the rule "Have some children", then our species would die out, but that hardly shows it is morally wrong not to

have any children. Luckily, our species will not die out if everyone is permitted not to have children, since enough people want to have children. Thus, instead of asking, "What would happen if everybody did that?", rule consequentialists should ask, "What would happen if everybody were permitted to do that?" People are permitted to do what violates no accepted rule, so asking what would happen if everybody were permitted to do an act is just the flip side of asking what would happen if people accepted a rule that forbids that act. Such acceptance rule consequentialists then claim that an act is morally wrong if and only if it violates a rule whose acceptance has better consequences than the acceptance of any incompatible rule. In some accounts, a rule is accepted when it is built into individual consciences (Brandt 1992). Other rule utilitarians, however, require that moral rules be publicly known (Gert 2005; cf. Sinnott-Armstrong 2003b) or built into public institutions (Rawls 1955). Then they hold what can be called public acceptance rule consequentialism: an act is morally wrong if and only if it violates a rule whose public acceptance maximizes the good.

The indirectness of such rule utilitarianism provides a way to remain consequentialist and yet capture the common moral intuition that it is immoral to perform the transplant in the above situation. Suppose people generally accepted a rule that allows a doctor to transplant organs from a healthy person without consent when the doctor believes that this transplant will maximize utility. Widely accepting this rule would lead to many transplants that do not maximize utility, since doctors (like most people) are prone to errors in predicting consequences and weighing utilities. Moreover, if the rule is publicly known, then patients will fear that they might be used as organ sources, so they would be less likely to go to a doctor when they need one. The medical profession depends on trust that this public rule would undermine. For such reasons, some rule utilitarians conclude that it would not maximize utility for people generally to accept a rule that allows doctors to transplant organs from unwilling donors. If this claim is correct, then rule utilitarianism implies that it is morally wrong for a particular doctor to use an unwilling donor, even for a particular transplant that would have better consequences than any alternative even from the doctor's own perspective. Common moral intuition is thereby preserved.

Rule utilitarianism faces several potential counterexamples (such as whether public rules allowing slavery could sometimes maximize utility) and needs to be formulated more precisely (particularly in order to avoid collapsing into act-utilitarianism; cf. Lyons 1965). Such details are discussed in another entry in this encyclopedia (see Hooker on rule-consequentialism). Here I just want to point out that direct consequentialists find it weird to judge a particular act by the consequences of something else (Smart 1956). Why should mistakes by other doctors in other cases make this doctor's act morally wrong, when this doctor knows for sure that he is not mistaken in this case? Rule consequentialists can respond that we should not claim special rights or permissions that we are not willing to grant to every other person, and that it is arrogant to think we are less prone to mistakes than other people are. However, this doctor can reply that he is willing to give everyone the right to violate the usual rules in the rare cases when they do know for sure that violating those rules really maximizes utility. Anyway, even if rule utilitarianism accords with some common substantive moral intuitions, it still seems counterintuitive in other ways. This makes it worthwhile to consider how direct consequentialists can bring their views in line with common moral intuitions, and whether they need to do so.

**6. Consequences for Whom? Limiting the Demands of Morality**
Another popular charge is that classic utilitarianism demands too much, because it requires us to do acts that are or should be moral options (neither obligatory nor forbidden). (Scheffler 1982) For example, imagine that my old shoes are serviceable but dirty, so I want a new pair of shoes that costs $100. I could wear my old shoes and give the $100 to a charity that will use my money to save someone else's life. It would seem to maximize utility for me to give the $100 to the charity. If it is morally wrong to do anything other than what maximizes utility, then it is morally wrong for me to buy the shoes. But buying the shoes does not seem morally wrong. It might be morally better to give the money to charity, but such contributions seem supererogatory, that is, above and beyond the call of duty. Of course, there are many more cases like this. When I watch television, I always (or almost always) could do more good by helping others, but it does not seem morally wrong to watch television. When I choose to teach philosophy rather than working for CARE or the Peace Corps, my choice probably fails to maximize utility overall. If we were required to maximize utility, then we would have to make very different choices in many areas of our lives. The requirement to maximize utility, thus, strikes many people as too demanding because it interferes with the personal decisions that most of us feel should be left up to the individual.

Some utilitarians respond by arguing that we really are morally required to change our lives so as to do a lot more to increase overall utility (see Kagan 1989, P. Singer 1993, and Unger 1996). Such hard-liners claim that

most of what most people do is morally wrong, because most people rarely maximize utility. Some such wrongdoing might be blameless when agents act from innocent or even desirable motives, but it is still supposed to be moral wrongdoing. Opponents of utilitarianism find this claim implausible, but it is not obvious that their counter-utilitarian intuitions are reliable or well-grounded (Murphy 2000, chs. 1–4; cf. Mulgan 2001).

Other utilitarians blunt the force of the demandingness objection by limiting direct utilitarianism to what people morally ought to do. Even if we morally ought to maximize utility, it need not be morally wrong to fail to maximize utility. John Stuart Mill, for example, argued that an act is morally wrong only when both it fails to maximize utility and its agent is liable to punishment for the failure (Mill 1861). It does not always maximize utility to punish people for failing to maximize utility (cf. Sinnott-Armstrong 2005). Thus, on this view, it is not always morally wrong to fail to do what one morally ought to do. If Mill is correct about this, then utilitarians can say that we ought to give much more to charity, but we are not required or obliged to do so, and failing to do so is not morally wrong.

Many utilitarians still want to avoid the claim that we morally ought to give so much to charity. One way around this claim uses a rule-utilitarian theory of what we morally ought to do. If it costs too much to internalize rules implying that we ought to give so much to charity, then, according to such rule-utilitarianism, it is not true that we ought to give so much to charity (Hooker 2000, ch. 8).

Another route follows an agent-relative theory of value. If there is more value in benefiting oneself or one's family and friends than there is disvalue in letting strangers die (without killing them), then spending resources on oneself or one's family and friends would maximize the good. A problem is that such consequentialism would seem to imply that we morally ought not to contribute those resources to charity, although such contributions seem at least permissible.

More personal leeway could also be allowed by deploying the legal notion of proximate causation. When a starving stranger would stay alive if and only if one contributed to a charity, contributing to the charity still need not be the proximate cause of the stranger's life, and failing to contribute need not be the proximate cause of his or her death. Thus, if an act is morally right when it includes the most net good in its proximate consequences, then it might not be morally wrong either to contribute to the charity or to fail to do so.

Yet another way to reach this conclusion is to give up maximization and to hold instead that we morally ought to do what creates enough utility. This position is often described as satisficing consequentialism (Slote 1984). According to satisficing consequentialism, it is not morally wrong to fail to contribute to a charity if one contributes enough to other charities and if the money or time that one could contribute does create enough good, so it is not just wasted. (For criticisms, see Bradley 2006.) A related position is progressive consequentialism, which holds that we morally ought to improve the world or make it better than it would be if we did nothing, but we don't have to improve it as much as we can (Elliot and Jamieson, 2009). Both satisficing and progressive consequentialism allow us to devote some of our time and money to personal projects that do not maximize overall good.

Opponents still object that all such consequentialist theories are misdirected. When I decide to visit a friend instead of working for a charity, I can know that my act is not immoral even if I have not calculated that the visit will create enough overall good or that it will improve the world. These critics hold that friendship requires us to do certain favors for friends without weighing our friends' welfare impartially against the welfare of strangers. Similarly, if I need to choose between saving my drowning wife and saving a drowning stranger, it would be "one thought too many" (Williams 1981) for me to calculate the consequences of each act. I morally should save my wife straightaway without calculating utilities.

In response, utilitarians can remind critics that the principle of utility is intended as only a criterion of right and not as a decision procedure, so utilitarianism does not imply that people ought to calculate utilities before acting (Railton 1984). Consequentialists can also allow the special perspective of a friend or spouse to be reflected in agent-relative value assessments (Sen 1982, Broome 1991, Portmore 2001, 2003) or probability assessments (Jackson 1991). It remains controversial, however, whether any form of consequentialism can adequately incorporate common moral intuitions about friendship.

## 7. Arguments for Consequentialism

Even if consequentialists can accommodate or explain away common moral intuitions, that might seem only to answer objections without yet giving any positive reason to accept consequentialism. However, most people begin with the presumption that we morally ought to make the world better when we can. The question then is only whether any moral constraints or moral options need to be added to the basic consequentialist factor in moral reasoning. (Kagan 1989, 1998) If no objection reveals any need for anything beyond consequences, then consequences alone seem to determine what is morally right or wrong, just as consequentialists claim.

This line of reasoning will not convince opponents who remain unsatisfied by consequentialist responses to objections. Moreover, even if consequentialists do respond adequately to every proposed objection, that would not show that consequentialism is correct or even defensible. It might face new problems that nobody has yet recognized. Even if every possible objection is refuted, we might have no reason to reject consequentialism but still no reason to accept it.

In case a positive reason is needed, consequentialists present a wide variety of arguments. One common move attacks opponents. If the only plausible options in moral theory lie on a certain list (say, Kantianism, contractarianism, virtue theory, pluralistic intuitionism, and consequentialism), then consequentialists can argue for their own theory by criticizing the others. This disjunctive syllogism or process of elimination will be only as strong as the set of objections to the alternatives, and the argument fails if even one competitor survives. Moreover, the argument assumes that the original list is complete. It is hard to see how that assumption could be justified.

Consequentialism also might be supported by an inference to the best explanation of our moral intuitions. This argument might surprise those who think of consequentialism as counterintuitive, but in fact consequentialists can explain many moral intuitions that trouble deontological theories. Moderate deontologists, for example, often judge that it is morally wrong to kill one person to save five but not morally wrong to kill one person to save a million. They never specify the line between what is morally wrong and what is not morally wrong, and it is hard to imagine any non-arbitrary way for deontologists to justify a cutoff point. In contrast, consequentialists can simply say that the line belongs wherever the benefits outweigh the costs (including any bad side effects). Similarly, when two promises conflict, it often seems clear which one we should keep, and that intuition can often be explained by the amount of harm that would be caused by breaking each promise. In contrast, deontologists are hard pressed to explain which promise is overriding if the reason to keep each promise is simply that it was made (Sinnott-Armstrong 2009). If consequentialists can better explain more common moral intuitions, then consequentialism might have more explanatory coherence overall, despite being counterintuitive in some cases. (Compare Sidgwick 1907, Book IV, Chap. III; and Sverdlik 2011.) And even if act consequentialists cannot argue in this way, it still might work for rule consequentialists (such as Hooker 2000).

Consequentialists also might be supported by deductive arguments from abstract moral intuitions. Sidgwick (1907, Book III, Chap. XIII) seemed to think that the principle of utility follows from very general principles of rationality and universalizability.

Other consequentialists are more skeptical about moral intuitions, so they seek foundations outside morality, either in non-normative facts or in non-moral norms. Mill (1861) is infamous for his "proof" of the principle of utility from empirical observations about what we desire (cf. Sayre-McCord 2001). In contrast, Hare (1963, 1981) tries to derive his version of utilitarianism from substantively neutral accounts of morality, of moral language, and of rationality.

Yet another argument for a kind of consequentialism is contractarian. Harsanyi (1977, 1978) argues that all informed, rational people whose impartiality is ensured because they do not know their place in society would favor a kind of consequentialism. Broome (1991) elaborates and extends Harsanyi's argument.

Other forms of arguments have also been invoked on behalf of consequentialism (e.g. Cummiskey 1996, P. Singer 1993; Sinnott-Armstrong 1992). However, each of these arguments has also been subjected to criticisms.

Even if none of these arguments proves consequentialism, there still might be no adequate reason to deny consequentialism. We might have no reason either to deny consequentialism or to assert it. Consequentialism

could then remain a live option even if it is not proven.

**Bibliography**

- Adams, R.M., 1976. "Motive Utilitarianism", Journal of Philosophy, 73: 467–81.
- Bales, R. E., 1971. "Act-utilitarianism: account of right-making characteristics or decision-making procedures?", American Philosophical Quarterly, 8: 257–65.
- Bayles, M. (ed.), 1968. Contemporary Consequentialism, Garden City, NY; Doubleday.
- Bennett, J., 1989. "Two Departures from Consequentialism", Ethics, 100: 54–66.
- ——, 1995. The Act Itself, New York: Oxford University Press.
- Bentham, J., 1843. Rationale of Reward, Book 3, Chapter 1, in The Works of Jeremy Bentham, J. Bowring (ed.), Edinburgh: William Tait.
- Bentham, J., 1961. An Introduction to the Principles of Morals and Legislation, Garden City: Doubleday. Originally published in 1789.
- Bradley, B., 2005. "Virtue Consequentialism", Utilitas, 17: 282–298.
- Bradley, B., 2006. "Against Satisficing Consequentialism", Utilitas, 18: 97–108.
- Brandt, R., 1979. A Theory of the Good and the Right, New York: Oxford University Press.
- ——, 1992. Morality, Utilitarianism, and Rights, Cambridge: Cambridge University Press.
- Brink, D., 1986. "Utilitarian Morality and the Personal Point of View", Journal of Philosophy, 83: 417–38.
- ——, 1989. Moral Realism and the Foundations of Ethics, New York: Cambridge University Press.
- ——, 2006. "Some Forms and Limits of Consequentialism", in The Oxford Handbook of Ethical Theory, D. Copp (ed.), Oxford: Clarendon Press.
- Broome, J., 1991. Weighing Goods, Oxford: Basil Blackwell.
- Brown, C., 2011. "Consequentialize This", Ethics, 121: 749–71.
- Carritt, E. F., 1947. Ethical and Political Thinking, Oxford: Oxford University Press.
- Chang, R., 1997. Incommensurability, Incomparability, and Practical Reason, Cambridge: Harvard University Press.
- Chappell, T., 2001. "Options Ranges", Journal of Applied Philosophy, 18 (2): 107–118.
- Coakley, M., 2015. "Interpersonal Comparisons of the Good: Epistemic Not Impossible", Utilitas, doi: 10.1017/S0953820815000266.
- Cummiskey, D., 1996. Kantian Consequentialism, New York: Oxford University Press.
- Darwall, S. (ed.), 2003. Consequentialism, Oxford: Blackwell.
- De Brigard, F., 1991. "If You Like It, Does It Matter If It's Real?", Philosophical Psychology, 23: 43–57.
- Dreier, J., 1993. "Structures of Normative Theories", Monist, 76: 22ff.
- ——, 2011. "In Defense of Consequentializing", in Oxford Studies in Normative Ethics, M. Timmons (ed.), Oxford: Oxford University Press, 97–119.
- Driver, J., 2001a. Uneasy Virtue, New York: Cambridge University Press.
- —— (ed.), 2001b. Character and Consequentialism, Special Issue of Utilitas, 13 (2).
- Feldman, F., 1986. Doing the Best We Can, Boston: D. Reidel.
- ——, 1997. Utilitarianism, Hedonism, and Desert, New York: Cambridge University Press.
- ——, 2004. Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism, New York: Oxford University Press.
- Foot, P., 1967. "Abortion and the Doctrine of Double Effect", Oxford Review, 5: 28–41.
- ——, 1983. "Utilitarianism and the Virtues", Proceedings and Addresses of the American Philosophical Association, 57 (2): 273–83. Revised in Mind, 94 (1985): 196–209.
- Frey, R. G. (ed.), 1984. Utility and Rights, Oxford: Basil Blackwell.
- Geach, P., 1956. "Good and Evil", Analysis, XVII (2): 33–42.
- Frey, R. G. (ed.), 1984. Utility and Rights, Oxford: Basil Blackwell.
- Gert, B., 2005. Morality: Its Nature and Justification, New York: Oxford University Press, revised edition.
- Goodin, R. E., 1995. Utilitarianism as a Public Philosophy, New York: Cambridge University Press.
- Griffin, J., 1986. Well-Being, Oxford: Clarendon Press.
- Hare, R. M., 1963. Freedom and Reason, London: Oxford University Press.
- ——, 1981. Moral Thinking, Oxford: Clarendon Press.
- Harsanyi, J. C., 1977. "Morality and the Theory of Rational Behavior", Social Research, 44 (4): 623–56. Reprinted in Sen and Williams 1982.

- ——, 1978. "Bayesian Decision Theory and Utilitarian Ethics", The American Economic Review, 68: 223–8.
- Hart, H. L. A., and Honoré, T., 1985. Causation in the Law, Second Edition. Oxford: Clarendon Press.
- Hawkins, J., forthcoming. "The Experience Machine and the Experience Requirement", The Routledge Handbook of the Philosophy of Well-Being.
- Hooker, B., 2000. Ideal Code, Real World, Oxford: Clarendon Press.
- Hooker, B., Mason, E., and Miller, D. E., 2000. Morality, Rules, and Consequences, Edinburgh: Edinburgh University Press.
- Howard-Snyder, F., 1994. "The Heart of Consequentialism", Philosophical Studies, 76: 107–29.
- ——, 1996. "A New Argument for Consequentialism? A Reply to Sinnott-Armstrong", Analysis, 56: 111–115.
- Hurka, T., 1993. Perfectionism, New York: Oxford University Press.
- ——, 2001. Virtue, Vice, and Value, New York: Oxford University Press.
- Hutcheson, F., 1965. A System of Moral Philosophy, Selection in Selby-Bigge (1965). Originally published in 1755.
- Jackson, F., 1991. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection", Ethics, 101: 461–82.
- Jamieson, D., 2005. "When Utilitarians Should be Virtue Theorists". Utilitas, 19 (02):160-.
- Jamieson, D., and Elliot, R., 2009. "Progressive Consequentialism". Philosophical Perspectives, 23: 241–251.
- Kagan, S., 1989. The Limits of Morality, Oxford: Clarendon Press.
- ——, 1998. Normative Ethics, Boulder: Westview.
- Kupperman, J. J., 1981. "A Case for Consequentialism", American Philosophical Quarterly, 18: 305–13.
- Lyons, D., 1965. Forms and Limits of Utilitarianism, Oxford: Clarendon Press.
- McCloskey, H. J., 1965. "A Non-Utilitarian Approach to Punishment", Inquiry, 8: 239–55.
- McNaughton, D., and Rawling, P., 1991. "Agent-Relativity and the Doing-Happening Distinction", Philosophical Studies, 63: 167–85.
- ——, 1992. "Honoring and Promoting Values", Ethics, 102: 835–43.
- Mill, J. S., 1861. Utilitarianism, edited with an introduction by Roger Crisp. New York: Oxford University Press, 1998.
- Moore, G. E., 1903. Principia Ethica, Cambridge: Cambridge University Press.
- ——, 1912. Ethics, New York: Oxford University Press.
- Mulgan, T., 2001. The Demands of Consequentialism, Oxford: Clarendon Press.
- Murphy, L., 2000. Moral Demands in Nonideal Theory, New York: Oxford University Press.
- Norcross, A., 1997. "Comparing Harms: Headaches and Human Lives", Philosophy and Public Affairs, 26: 135–67.
- Nozick, R., 1974. Anarchy, State, and Utopia, New York: Basic Books.
- Nussbaum, Martha C., 2000. Women and Human Development: The Capabilities Approach, New York: Cambridge University Press.
- Parfit, D., 1984. Reasons and Persons, Oxford: Clarendon Press.
- Pettit, P., 1984. "Satisficing Consequentialism", Proceedings of the Aristotelian Society 58: 165–76.
- —— (ed.), 1993. Consequentialism, Aldershot: Dartmouth.
- ——, 1997. "The Consequentialist Perspective" in Three Methods of Ethics, by M. Baron, P. Pettit, and M. Slote. Oxford: Blackwell.
- Pettit, P., and Brennan, G., 1986. "Restrictive Consequentialism", Australasian Journal of Philosophy, 64: 438–55.
- Pettit, P., and Smith, M., 2000. "Global Consequentialism" in Hooker et al, pp. 121–33.
- Plato, Philebus, trans. D. Frede, Indianapolis: Hackett Publishing, 1993.
- Portmore, Douglas W., 2001. "Can an Act-Consequentialist Theory be Agent-Relative?" American Philosophical Quarterly, 38: 363–77.
- ——, 2003. "Position-Relative Consequentialism. Agent-Centered Options, and Supererogation", Ethics, 113: 303–32.
- ——, 2009. "Consequentializing", Philosophy Compass, 4.
- Railton, P., 1984. "Alienation, Consequentialism, and the Demands of Morality", Philosophy and Public Affairs, 13: 134–71. Reprinted in Railton 2003.

- ——, 2003. Facts, Values, and Norms: Essays toward a Morality of Consequence, Cambridge: Cambridge University Press.
- Rawls, J., 1955. "Two Concepts of Rules", Philosophical Review, 64: 3–32.
- ——, 1971. A Theory of Justice, Cambridge: Harvard University Press.
- Regan, D., 1980. Utilitarianism and Cooperation, Oxford: Clarendon Press.
- Roberts, M. A., 2002. "A New Way of Doing the Best That We Can: Person-Based Consequentialism and the Equality Problem", Ethics, 112, 2: 315–50.
- Ross, W. D., 1930. The Right and the Good, Oxford: Clarendon Press.
- Sayre-McCord, G., 2001. "Mill's 'Proof' of the Principle of Utility: A More than Half-Hearted Defense", in Moral Knowledge, E. F. Paul, F. D. Miller, and J. Paul (eds.), New York: Cambridge University Press, pp. 330–60.
- Scanlon, T. M., 1982. "Contractualism and Utilitarianism", in Sen and Williams (eds.) 1982.
- Scarre, G., 1996. Utilitarianism, London: Routledge.
- Scheffler, S., 1982. The Rejection of Consequentialism, Oxford: Clarendon Press. Revised edition 1994.
- —— (ed.), 1988. Consequentialism and Its Critics, Oxford: Oxford University Press.
- Schneewind, Jerome, 1990. Moral Philosophy from Montaigne to Kant, New York: Cambridge University Press, 2002.
- Selby-Bigge, L. A. (ed.), 1965. British Moralists, New York: Dover.
- Sen, A., 1979. "Utilitarianism and Welfarism", Journal of Philosophy, 76: 463–89.
- ——, 1982. "Rights and Agency", Philosophy and Public Affairs, 11 (1): 3–39.
- ——, 1985. "Well-Being, Agency, and Freedom", Journal of Philosophy, 82 (4): 169–221.
- ——. 2002. Rationality and Freedom, Cambridge: Harvard University Press.
- Sen, A., and Williams, B. (eds.), 1982. Utilitarianism and Beyond, Cambridge: Cambridge University Press.
- Shaw, W. H., 1999. Contemporary Ethics: Taking Account of Utilitarianism, Malden: Blackwell.
- Sidgwick, H., 1907. The Methods of Ethics, Seventh Edition. London: Macmillan. First Edition 1874.
- Singer, M., 1977. "Actual Consequence Utilitarianism", Mind, 86: 67–77.
- Singer, P., 1974. "Sidgwick and Reflective Equilibrium", Monist, 58: 490–517.
- ——, 1993. Practical Ethics, Second Edition. Cambridge: Cambridge University Press.
- Sinnott-Armstrong, W., 1988. Moral Dilemmas, Oxford: Blackwell.
- ——, 1992. "An Argument for Consequentialism", Philosophical Perspectives, 6: 399–421.
- ——, 2001. "What is Consequentialism? A Reply to Howard-Snyder", Utilitas, 13 (3): 342-9.
- ——, 2003a. "For Goodness' Sake", Southern Journal of Philosophy, 41 (Supplement): 83–91.
- ——, 2003b. "Gert Contra Consequentialism" in Rationality, Rules, and Ideals, W. Sinnott-Armstrong and R. Audi (eds.), New York: Rowman and Littlefield.
- ——, 2005. "You Ought to be Ashamed of Yourself (When you Violate an Imperfect Moral Obligation)", Philosophical Issues, 15: 193–208.
- ——, 2009. "How strong is this obligation? An argument for consequentialism from concomitant variation", Analysis, 69: 438–442.
- Skorupski, J., 1995. "Agent-Neutrality, Consequentialism, Utilitarianism … A Terminological Note," Utilitas, 7: 49–54.
- ——, 1999. Ethical Explorations, Oxford: Oxford University Press.
- Slote, M., 1984. "Satisficing Consequentialism", Proceedings of the Aristotelian Society, 58: 139–63.
- ——, 1985. Common-Sense Morality and Consequentialism, London: Routledge and Kegan Paul.
- Smart, J. J. C., 1956. "Extreme and Restricted Utilitarianism", The Philosophical Quarterly, 6: 344–54.
- ——, 1973. "An Outline of a System of Utilitarian Ethics" in Utilitarianism: For and Against, by J.J.C. Smart and B. Williams. Cambridge: Cambridge University Press, pp. 3–74.
- Smart, R. N., 1958. "Negative Utilitarianism", Mind, 67: 542–3.
- Sosa, D., 1993. "Consequences of Consequentialism", Mind, 102 (405): 101–22.
- Sprigge, T. L. S., 1965. "A Utilitarian Reply to Dr. McCloskey", Inquiry, 8: 264–91.
- ——, 1988. The Rational Foundations of Ethics, London: Routledge & Kegan Paul.
- Sumner, L. W., 1987. The Moral Foundations of Rights, Oxford: Clarendon Press.
- ——, 1996. Welfare, Happiness, and Ethics, Oxford: Clarendon Press.
- Sverdlik, Steven, 2011. Motives and Rightness, Oxford: Oxford University Press.
- Tannsjo, Torbjorn, 1998. Hedonistic Utilitarianism, Edinburgh: Edinburgh University Press.

- Thomson, J. J., 1976. "Killing, Letting Die, and the Trolley Problem", The Monist, 59: 204–17.
- ——, 1994. "Goodness and Utilitarianism", Proceedings and Addresses of the American Philosophical Association, 67 (4): 7–21.
- ——, 2001. Goodness and Advice, Amy Gutmann (ed.), Princeton; Princeton University Press.
- Unger, P., 1996. Living High and Letting Die, New York: Oxford University Press.
- Williams, B., 1973. "A Critique of Utilitarianism" in Utilitarianism: For and Against, by J.J.C. Smart and B. Williams. Cambridge: Cambridge University Press, pp. 77–150.
- ——, 1981. "Persons, Character, and Morality", in B. Williams, Moral Luck, Cambridge: Cambridge University Press.