# Introduction to Utilitarianism

> *"The utilitarian doctrine is, that happiness is desirable, and the only thing desirable, as an end; all other things being only desirable as means to that end."*
>
> - John Stuart Mill [1]

## Table of Contents


1. Why Do We Need Moral Theories?
2. What Is Utilitarianism?
   - Classical Utilitarianism
3. Utilitarianism and Practical Ethics
4. Acting on Utilitarianism
5. Arguments for Utilitarianism
   - What Fundamentally Matters
   - The Veil of Ignorance
   - Avoiding Status Quo Bias
6. Track Record
7. Objections to Utilitarianism
   - The Rights Objection
   - The Demandingness Objection
8. Conclusion
9. Resources and Further Reading


Utilitarianism was developed to answer the question of what we morally ought to do, and why. Its core idea is that we ought to act to improve the well-being of everyone by as much as possible. Compared to other ethical theories, it is unusually demanding and may tell us to make substantial changes to how we lead our lives. Perhaps more so than any other ethical theory, it has produced a fierce philosophical debate between its proponents and its critics.

## Why Do We Need Moral Theories?

When we make moral judgments in everyday life, we often rely on our intuition. If you ask yourself whether or not it is wrong to eat meat, or to lie to a friend, or to buy sweatshop goods, you probably have a strong gut moral view on the topic. But there are problems with relying merely on our moral intuition.

Historically, people held beliefs we now consider morally horrific. In Western societies, it was once firmly believed to be intuitively obvious that people of color and women have fewer rights than white men; that homosexuality is wrong; and that it was permissible to own slaves. We now see these moral intuitions as badly misguided. This historical track record gives us reason to be concerned that we, in the modern era, may also be unknowingly responsible for serious, large-scale wrongdoing. It would be a very lucky coincidence if the present generation were the first generation whose intuitions were perfectly morally correct. [2]

Also, people have conflicting moral intuitions, and we need a way to resolve these disagreements. The project of moral philosophy is to reflect on our competing moral intuitions and develop a theory that will tell us which actions are right or wrong, and why. This will then allow us to identify which moral judgments of today are misguided, enabling us to make moral progress and act more ethically.

One of the most prominent and influential attempts to create such a theory is *utilitarianism*. Utilitarianism was developed by the philosophers Jeremy Bentham and John Stuart Mill, who drew on ideas going back to the ancient Greeks. Their utilitarian views have been widely discussed since they first espoused them, and have had significant influence in economics and public policy.

## What Is Utilitarianism?

The core idea of utilitarianism is that we ought to act to improve the well-being of everyone by as much as possible.

A more precise definition of utilitarianism is as follows:

> Utilitarianism is the view that one ought always to promote the sum total of well-being. [3]

Sometimes philosophers talk about "welfare" or "utility" rather than "well-being", but these words are typically used to mean the same thing. [4] Utilitarianism is most commonly applied to evaluate the rightness of actions, but the theory can also evaluate other things, like rules, policies, motives, virtues, and social institutions. It is perhaps unfortunate that the clinical-sounding term "utilitarianism" caught on as a name, especially since in common speech the word "utilitarian" is easily confused with joyless functionality or even outright selfishness.

All ethical theories belonging to the utilitarian family share four defining elements: (i) consequentialism, (ii) welfarism, (iii) impartiality, and (iv) aggregationism.

- **Consequentialism** is the view that one ought always to promote good outcomes.

- **Welfarism** is the view that only the *welfare* (also called *well-being*) of individuals determines the value of an outcome.
- **Impartiality** is the view that the identity of individuals is irrelevant to the value of an outcome. Utilitarians hold, more specifically, that equal weight must be given to the interests of all individuals.
- **Aggregationism** is the view that the value of the world is the sum of the values of its parts, where these parts are local phenomena such as experiences, lives, or societies. [5]

Utilitarianism's rivals are theories that deny one or more of the above four elements. For example, they might hold that actions can be inherently right or wrong regardless of their consequences, or that some outcomes are good even if they do not increase the welfare of any individual, or that morality allows us to be partial towards our friends and families.

We cover the four elements of utilitarianism and its variants in greater depth in Chapter 2: Elements and Types of Utilitarianism.

### Classical Utilitarianism

The early utilitarians—Jeremy Bentham, John Stuart Mill, and Henry Sidgwick—were *classical utilitarians*. Classical utilitarianism is distinct from other utilitarian theories in that it accepts these two additional elements: First, it accepts hedonism as a theory of well-being.

- **Hedonism** is the view that well-being consists in, and only in, the balance of positive over negative conscious experiences. For readability, we will call positive conscious experiences *happiness* and negative conscious experiences *suffering*.

Second, classical utilitarianism accepts the total view of population ethics.

- **The total view of population ethics** regards one outcome as better than another if and only if it contains greater total well-being.

Classical utilitarianism can be defined as follows:

> **Classical utilitarianism is the view that one ought always to promote the sum total of happiness over suffering.**

## Utilitarianism and Practical Ethics

Utilitarianism is a demanding ethical theory that may require us to substantially change how we act. Utilitarianism says that we should make helping others a very significant part of our lives. In helping others, we should try to use our resources to do the most good, impartially considered, that we can.

According to utilitarianism, we should extend our moral concern to all sentient beings, meaning every individual capable of experiencing positive or negative conscious states. On this basis, a priority for utilitarians may be to help society to continue to widen its moral circle of concern. For instance, we may want to persuade people they should help not just those in their own country, but also people on the other side of the world; not just those of their own species but all sentient creatures; and not just people currently alive but any people whose lives they can affect.

Despite having a radically different approach to ethics than commonsense morality, utilitarianism generally endorses commonsense prohibitions. For practical purposes, the best course of action for a utilitarian is to try to do as much good as possible whilst still acting in accordance with commonsense moral virtues—like integrity, trustworthiness, law-abidingness, and fairness.

We discuss the implications of utilitarianism for moral deliberation in Chapter 6: Utilitarianism and Practical Ethics.

## Acting on Utilitarianism

There are many problems in the world today, some of which are extremely large in scale. Unfortunately, our resources are scarce, so as individuals and even as a global society we cannot solve all the world's problems at once. This means we must decide how to prioritize the resources we have. Not all ways of helping others are equally effective. By the lights of utilitarianism, we should choose carefully which moral problems to work on and by what means, based on where we can do the most good. This involves taking seriously the question of how we can best use our time and money to help others. Once again, utilitarianism urges us to consider the well-being of all individuals regardless of what species they belong to, what country they live in, and at what point in time they exist. With this in mind, a few moral problems appear especially pressing:

- *Global Health and Development*. Those in affluent countries are typically one hundred times richer than the poorest seven hundred million people in the world. We can radically improve the lives of the extreme poor, such as by providing basic medical care, at very little cost.[6]
- *Factory Farming*. Tens of billions of non-human animals are kept in horrific conditions in factory farms, undergoing immense unnecessary suffering. We could radically decrease this suffering at very little cost to society.
- *Existential Risks*. There will be vast numbers of people in the future, and their lives could be very good. Yet technological progress brings risks, such as from climate change, nuclear war, synthetic biology and artificial intelligence, that could endanger humanity's future. But if we can successfully navigate these risks, we can ensure a flourishing world for trillions of people yet to come.[7] There are three key means of helping those affected by the above moral concerns: donating money to effective charities, working in an impactful career, and convincing other people to do the same. For example, donations to the most effective global

health charities are expected to save a human life for under $5,000; this money may go even further when donated to address factory farming or existential risks. Choosing which career to pursue may be even more important again, since some careers allow us to do far more good than others.[8]

In the article Acting on Utilitarianism, we discuss what utilitarianism means for how to live ethically.

## Arguments for Utilitarianism

No one can prove that utilitarianism is the correct moral theory. Nor can anyone prove it for any of utilitarianism's rivals. What we can do is to consider the arguments for and against all plausible ethical theories, ultimately coming to an all-things-considered judgement about which is most compelling. In what follows, we will briefly outline three arguments for utilitarianism. We discuss these and several other arguments in more detail in Chapter 3: Arguments for Utilitarianism.

### What Fundamentally Matters

Moral theories serve to specify *what fundamentally matters*, and utilitarianism seems to offer a particularly compelling answer to this question.

Almost anyone would agree with utilitarianism that suffering is bad, and happiness is good. What could be more obvious? If anything matters morally, human well-being surely does. And it would be arbitrary to limit moral concern to our own species, so we should instead conclude that well-being generally is what matters. That is, we ought to want the lives of sentient beings to go as well as possible (whether that ultimately comes down to maximizing happiness, desire satisfaction, or other welfare goods).

Commonsense moral rules, such as those prohibiting theft and promise-breaking, can plausibly be given a utilitarian basis, as such rules generally promote overall well-being. If they did not, it would be hard to see what reason we would have to still want people to follow them. To follow *harmful* moral rules would seem like a kind of "rule worship", and not truly ethical at all.[9]

### The Veil of Ignorance

The second argument for utilitarianism is presented by the Nobel Prize-winning economist John Harsanyi.[10] Harsanyi studied situations where individuals had to make decisions while facing uncertainty. In his work he suggested that morality is about taking the perspective of society as a whole, that the right moral view is the one which you would have chosen if you did not know who you were going to be in society, and that utilitarianism is that view.

More precisely, imagine you had to decide how to structure society from behind a *veil of ignorance*. Behind this veil of ignorance, you know all the facts about each person's circumstances in society—

what their income is, how happy they are, how they are affected by social policies, and their preferences and likes. However, what you do not know is which of these people you are. You only know that you have an *equal chance of being any* of these people. Imagine, now, that you are trying to act in a rational and self-interested way—you are just trying to do whatever is best for yourself. How would you structure society?

Harsanyi proved that in this situation you will structure society to promote the sum total of everyone's well-being.[11] In other words, if you are rational and acting in self-interest and were put behind the veil of ignorance, you would come to use some version of utilitarianism as the principle to decide about the structure and rules of society.

### Avoiding Status Quo Bias

Opposition to utilitarian trade-offs—that is, benefiting some at a lesser cost to others—amounts to a kind of status quo bias, prioritizing the *preservation of privilege* over promoting well-being more generally.

Such conservatism might stem from the Just World fallacy: the mistake of assuming that the status quo is just, and that people naturally get what they deserve. Of course, reality offers no such guarantees of justice. What circumstances one is born into depends on sheer luck, including one's endowment of physical and cognitive abilities which may pave the way for future success or failure. Thus, even later in life we never manage to fully wrest back control from the whimsies of fortune and, consequently, some people are vastly better off than others despite being no more deserving. In such cases, why should we not be willing to benefit one person at a lesser cost to privileged others? They have no special entitlement to the extra well-being that fortune has granted them.[12]

Clearly, it is good for people to be well-off, and we certainly would not want to harm anyone unnecessarily.[13] However, if we can increase overall well-being by benefiting one person at the lesser cost to another, we should not refrain from doing so merely due to a prejudice in favor of the existing distribution.[14] It is easy to see why traditional elites would want to promote a "morality" which favors their entrenched interests. It is less clear why others should go along with such a distorted view of what (and who) matters.

See Chapter 3 for further discussion of the arguments for utilitarianism.

## Track Record

While not constituting an *argument* per se, it is worth noting that utilitarian moral reasoning has a strong track record of contributing to humanity's collective moral progress. The classical utilitarians of the 18th and 19th centuries had social and political attitudes that were far ahead of their time: As a progressive social reformer, Jeremy Bentham defended issues such as the separation of church and state; the abolition of slavery and of capital punishment; legal regulations

to protect criminals and non-human animals from cruel treatment[15]; and the decriminalization of homosexuality.[16] Indeed, his manuscripts on homosexuality were so liberal that his editor hid them from the public after Bentham's death. They were only published two centuries later.

John Stuart Mill defended the provision of social welfare for the poor and of freedom of speech. He was the second MP in the UK Parliament to call for women's suffrage[17] and advocated for gender equality more generally. In his essay *The Subjection of Women*[18] Mill argued that

> the principle which regulates the existing social relations between the two sexes — the legal subordination of one sex to the other — is wrong itself, and now one of the chief hindrances to human improvement; and that it ought to be replaced by a principle of perfect equality, admitting no power or privilege on the one side, nor disability on the other.[19]

In a similar vein, Henry Sidgwick advocated for women's education and the freedom of education from religious doctrines. Modern utilitarians like Peter Singer are outspoken advocates against pressing moral problems such as extreme poverty and factory farming.[20]

While the early proponents of utilitarianism were still far from getting everything right, their utilitarian reasoning led them to escape many of the moral prejudices of their time and reach more enlightened moral and political positions. Those of us living today are, of course, no less fallible than our forebears. To help overcome our own biases, our moral and political views may similarly benefit from being checked against utilitarian principles.

## Objections to Utilitarianism

Critics have raised many objections to utilitarianism, to which advocates of the theory have responded in turn.

Often, critics allege that utilitarianism runs counter to our commonsense moral intuitions. For example, according to utilitarianism, we should give the same moral consideration to distant strangers as to our friends and family. Many would regard this as extremely counterintuitive.

Utilitarian philosophers often respond to such concerns by arguing that many of our moral intuitions, including our counter-utilitarian intuitions, are significantly misguided. They are therefore happy to considerably revise those moral intuitions in light of theoretical considerations. They hold that while intuitions about particular cases should be given some weight in our moral deliberations—so, all else being equal, a moral theory is more plausible if it fits better with our intuitions—they are only one component of the assessment of a moral theory.

In what follows, we will state two of the most prominent critiques of utilitarianism. We will thoroughly discuss these and other objections in Chapter 8: Objections to Utilitarianism and Responses.

### The Rights Objection

According to commonsense morality and many non-utilitarian theories, there are certain *moral constraints* you should never, or rarely, violate. These constraints are expressed in moral rules like "do not lie!" and "do not kill!". These rules are intuitively very plausible. This presents a problem for utilitarianism. The reason for this is that utilitarianism not only specifies which outcomes are best—those having the highest overall level of well-being—but also says that it would be wrong to fail to realize these outcomes.

Sometimes, realizing the best outcome may require violating moral constraints against harming others—that is, violating their rights. For example, suppose there were five people waiting for an organ transplant and that you could save their lives if you killed one other person to harvest their organs. Intuitively, we would regard this as wrong, but it seems that utilitarianism would regard this as morally required. In general, since there is no reason to expect commonsense moral rules to always coincide with the best ways to act according to utilitarianism, we should think it likely for them to conflict sometimes.

### The Demandingness Objection

Many critics argue that utilitarianism is too demanding, because it requires us to always act such as to bring about the best outcome. The theory leaves no room for actions that are permissible yet do not bring about the best consequences; this is why some critics claim that utilitarianism is a morality only for saints.[21]

As a matter of fact, very few people, including utilitarian philosophers, live their life in perfect accordance with utilitarianism. For instance, consider that the money a person spends on dining out could pay for several bednets, each protecting two children in a low-income country from malaria for about two years.[22] From a utilitarian perspective, the benefit to the person from dining out is much smaller than the benefit to the children from not having malaria, so it would seem the person has acted wrongly in choosing to have a meal out. Analogous reasoning applies to how we use our time: the hours someone spends on social media should apparently be spent volunteering for a charity or working harder at one's job to earn more money to donate.

To many people, these extreme obligations of utilitarianism seem absurd at first glance. According to commonsense morality, we are permitted to spend most of our income on ourselves, our loved ones, and on our personal projects. Insofar as charity goes, commonsense morality holds that while it is good and praiseworthy to donate, it is not obligatory.

## Conclusion

What matters most for utilitarianism is bringing about the best consequences for the world. This involves improving the well-being of all individuals, regardless of their gender, race, species, and

their geographical or temporal location. Against this background, three key concerns for utilitarianism are helping the global poor, improving farmed animal welfare, and ensuring that the future goes well over the long term. Furthermore, utilitarianism is unusually demanding: it may require us to make benefiting others the main focus of our lives.

All utilitarian theories share the four elements of consequentialism, welfarism, impartiality, and aggregationism. The original and most influential version of utilitarianism is classical utilitarianism, which encompasses two further characteristics: hedonism and totalism. Hedonism is the view that well-being consists entirely of conscious experiences, such as happiness or suffering. The total view regards one outcome as better than another if and only if it contains a greater sum total of well-being.

Utilitarian ethics has an intellectual tradition spanning centuries, during which it has been cause for many heated debates among moral philosophers. On the one hand, critics of utilitarianism accuse the theory of disregarding moral rights and for being overly demanding, among other objections. On the other hand, advocates of utilitarianism have argued that the theory has attractive theoretical virtues such as avoiding status quo bias and offering a compelling account of what fundamentally matters. In addition, they make the case that if we had to decide about the structure of society behind a veil of ignorance, it would be rational for everyone to make their choices based on utilitarian principles.

The next chapter discusses the four elements of utilitarian theories in greater depth and introduces several variants of utilitarianism.

Next Chapter: Elements and Types of Utilitarianism

## How to Cite This Page

MacAskill, W., Meissner, D., and Chappell, R.Y. (2022). Introduction to Utilitarianism. In R.Y. Chappell, D. Meissner, and W. MacAskill (eds.), *An Introduction to Utilitarianism* https://www.utilitarianism.net/introduction-to-utilitarianism, accessed 2/12/2023.

## Resources and Further Reading

### Introduction

- Utilitarianism: Crash Course Philosophy #36

- Peter Singer & Katarzyna de Lazari-Radek (2017). *Utilitarianism: A Very Short Introduction*. Oxford: Oxford University Press.
- Krister Bykvist (2010). *Utilitarianism: A Guide for the Perplexed*. London: Continuum.

## The Classics

- Jeremy Bentham (1789). An Introduction to the Principles of Morals and Legislation.
- John Stuart Mill (1863). Utilitarianism.
- Henry Sidgwick (1874). The Methods of Ethics.

## Further Reading

- Julia Driver (2014). The History of Utilitarianism. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.).
- Bart Schultz (2017). *The Happiness Philosophers: The Lives and Works of the Great Utilitarians*. Princeton: Princeton University Press.
- James Crimmins (2017). *The Bloomsbury Encyclopedia of Utilitarianism*. Bloomsbury.
- Derek Parfit (2011/17). *On What Matters*. Oxford: Oxford University Press.
- Yew-Kwang Ng (1990). Welfarism and Utilitarianism: A Rehabilitation. *Utilitas*. 2(2): 171–193.

---

1. Mill, J. S. (1863). *Utilitarianism*. Kitchener: Batoche Books., p. 35 ↩
2. For more details, see Williams, E. G. (2015). The Possibility of an Ongoing Moral Catastrophe. *Ethical Theory and Moral Practice*, *18*(5), 971–982. doi: 10.1007/s10677-015-9567-7 ↩
3. This definition applies to a fixed-population setting, where one's actions do not affect the number or identity of people. There are utilitarian theories that differ in how they deal with variable-population settings. This is a technical issue, relevant to the discussion of population ethics. ↩
4. However, when economists use the term "utility" they typically refer, instead, to the numerical representation of an individual's preferences. ↩
5. This definition applies to a fixed-population setting, where one's actions do not affect the number or identity of people. There are aggregationist theories that differ in how they deal with variable-population settings. This is a technical issue, relevant to the discussion of population ethics. ↩
6. For a detailed discussion of the ethics of Global Health and Development, we recommend Peter Singer's book The Life You Can Save (available for free download). ↩

7. For a detailed discussion of existential risks, we recommend Toby Ord's book *The Precipice: Existential Risk and the Future of Humanity*. ↩

8. Note that Professor William MacAskill, coauthor of this website, is a cofounder of 80,000 Hours. ↩

9. Smart, J.J.C. (1956). Extreme and restricted utilitarianism. *The Philosophical Quarterly*, 6 (25): 344 - 354. ↩

10. This argument was originally proposed by Harsanyi, though nowadays it is more often associated with John Rawls, who arrived at a different conclusion. ↩

11. Harsanyi formalized his argument for utilitarianism in Harsanyi, J. (1978). Bayesian Decision Theory and Utilitarian Ethics. *The American Economic Review, 68*(2), 223–228. For discussion about his proof, see Greaves, H. (2017). A Reconsideration of the Harsanyi–Sen–Weymark Debate on Utilitarianism. *Utilitas, 29*(2), 175–213. ↩

12. In a similar vein, Derek Parfit wrote that "Some of us ask how much of our wealth we rich people ought to give to these poorest people. But that question wrongly assumes that our wealth is ours to give. This wealth is legally ours. But these poorest people have much stronger moral claims to some of this wealth. We ought to transfer to these people (...) at least ten per cent of what we earn". Parfit, D. (2017). *On What Matters, Volume Three*. Oxford University Press., pp. 436-37 ↩

13. On the topic of sacrifice, John Stuart Mill wrote that "The utilitarian morality does recognise in human beings the power of sacrificing their own greatest good for the good of others. It only refuses to admit that the sacrifice is itself a good. A sacrifice which does not increase, or tend to increase, the sum total of happiness, it considers as wasted." Mill, J. S. (1863). Chapter 2: What Utilitarianism Is, Utilitarianism. ↩

14. However, this does not mean that utilitarianism will strive for perfect equality in material outcomes or even well-being. Joshua Greene notes that "a world in which everyone gets the same outcome no matter what they do is an idle world in which people have little incentive to do anything. Thus, the way to maximize happiness is not to decree that everyone gets to be equally happy, but to encourage people to behave in ways that maximize happiness. When we measure our moral success, we count everyone's happiness equally, but achieving success almost certainly involves inequality of both material wealth and happiness. Such inequality is not ideal, but it's justified on the grounds that, without it, things would be worse overall. Greene, J. (2013). *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them* . The Penguin Press., p. 163; see also: The Equality Objection to Utilitarianism. ↩

15. For instance, Bentham commented on the issue of animal protection: "*the question is not, Can they reason? nor, Can they talk? but, Can they suffer? Why should the law refuse its protection*

*to any sensitive being? The time will come when humanity will extend its mantle over everything which breathes. We have begun by attending to the condition of slaves; we shall finish by softening that of all the animals which assist our labors or supply our wants.*" Bentham, J. (1789). An Introduction to the Principles of Morals and Legislation. Bennett, J. (ed.)., pp. 143-144 ↩

16. Cf. Bentham, J. (1789). An Introduction to the Principles of Morals and Legislation. Bennett, J. (ed.).; and Campos Boralevi, L. (2012).Bentham and the Oppressed. Berlin: De Gruyter. ↩

17. The UK Parliament, The 1866 Women's Suffrage petition: the first mass Votes for Women petition. ↩

18. Mill attributes many of the ideas in *The Subjection of Women* to his wife, Harriet Taylor Mill. See Mill, J. S. (1873). *Autobiography*. Bennet, J. (ed.)., p. 166 ↩

19. Mill, J. S. (1869). *The Subjection of Women*. Bennet, J. (ed.)., p. 1 ↩

20. On extreme poverty see: Singer, P. (2009). *The Life You Can Save: Acting Now to End World Poverty*. Bainbridge Island, WA: Penguin Random House LLC.

    On factory farming see: Singer, P. (1975). *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York: HarperCollins. ↩

21. Cf. Wolf, S. (1982). Moral Saints. *The Journal of Philosophy*, *79*(8), 419. doi: 10.2307/2026228 ↩

22. GiveWell (2019). Against Malaria Foundation. ↩