[Template:Hatnote](/wiki/Template:Hatnote" \o "Template:Hatnote) [thumb|190px|The symbol sigma is used in mathematics to represent](/wiki/File:Greek_uc_sigma.svg) [summation](/wiki/Summation), a process similarly done in [total utilitarianism](/wiki/Average_and_total_utilitarianism). [Template:Utilitarianism](/wiki/Template:Utilitarianism)

**Utilitarianism** is a theory in [normative ethics](/wiki/Normative_ethics) holding that the best moral action is the one that maximizes [utility](/wiki/Utility). Utility is defined in various ways, but is usually related to the [well-being](/wiki/Well-being) of [sentient](/wiki/Sentience) entities. Originally, [Jeremy Bentham](/wiki/Jeremy_Bentham), the founder of Utilitarianism, defined utility as the aggregate pleasure after deducting suffering of all involved in any action. [John Stuart Mill](/wiki/John_Stuart_Mill) expanded this concept of utility to include not only the quantity, but quality of pleasure, while focusing on rules, instead of individual moral actions. Maximizes the profit, benefit or pleasure, while maintaining the lowest level of pain and loss. Others have rejected that pleasure has positive value and have advocated negative utilitarianism, which defines utility only in terms of suffering. As opposed to this [hedonistic](/wiki/Hedonism) view, some define utility with relation to [preference](/wiki/Preference) satisfaction whereas others believe that a range of values can be included in its definition.

Utilitarianism is a form of [consequentialism](/wiki/Consequentialism), which states that the consequences of any action are the only standard of right and wrong. This view can be contrasted or combined with [virtue ethics](/wiki/Virtue_ethics) which holds virtue as a [moral good](/wiki/Value_(ethics)). Some believe that one's intentions are also ethically important. Utilitarianism is distinctly different from other forms of consequentialism such as [egoism](/wiki/Ethical_egoism) as it considers [all interests equally](/wiki/Equal_consideration_of_interests). Proponents of utilitarianism have been split about whether individual acts should conform to utility ([act utilitarianism](/wiki/Act_utilitarianism)) or whether [agents](/wiki/Agency_(philosophy)) should conform to ethical rules ([rule utilitarianism](/wiki/Rule_utilitarianism)). Utilitarians additionally remain split about whether utility should be calculated as an aggregate ([total utilitarianism](/wiki/Average_and_total_utilitarianism)) or an average (average utilitarianism).

Historically, hedonism can be traced back to [Aristippus](/wiki/Aristippus) and [Epicurus](/wiki/Epicurus), who viewed happiness as the only good. Bentham is, however, credited with founding utilitarianism when he wrote *An Introduction to the Principles of Morals and Legislation.* Since Bentham, prominent utilitarians have included John Stuart Mill, [Henry Sidgwick](/wiki/Henry_Sidgwick), [R. M. Hare](/wiki/R._M._Hare) and [Peter Singer](/wiki/Peter_Singer). The philosophy has been applied to modern issues including the suffering of non-human animals. Specifically, utilitarianism has been applied to the [ethics of raising animals for food](/wiki/Ethics_of_eating_meat).

Opponents of utilitarianism have criticized it for many reasons. Some have said that utilitarianism ignores justice while others contend that utilitarianism is impractical. Specific criticisms have included the [mere addition paradox](/wiki/Mere_addition_paradox) and the [utility monster](/wiki/Utility_monster). Others have said that pleasure is not commensurable across people with varying identities and thus the idea of aggregating utility is impossible.

## Contents

* 1 Etymology[[edit](/index.php?title=(none)&action=edit&section=1)]
* 2 Historical background[[edit](/index.php?title=(none)&action=edit&section=2)]
* 3 Classical utilitarianism[[edit](/index.php?title=(none)&action=edit&section=3)]
  + 3.1 Jeremy Bentham[[edit](/index.php?title=(none)&action=edit&section=4)]
  + 3.2 John Stuart Mill[[edit](/index.php?title=(none)&action=edit&section=5)]
    - 3.2.1 Higher and lower pleasures[[edit](/index.php?title=(none)&action=edit&section=6)]
    - 3.2.2 Mill's "proof" of the principle of utility[[edit](/index.php?title=(none)&action=edit&section=7)]
* 4 Twentieth-century developments[[edit](/index.php?title=(none)&action=edit&section=8)]
  + 4.1 Ideal utilitarianism[[edit](/index.php?title=(none)&action=edit&section=9)]
  + 4.2 Act and rule utilitarianism[[edit](/index.php?title=(none)&action=edit&section=10)]
  + 4.3 Two-level utilitarianism[[edit](/index.php?title=(none)&action=edit&section=11)]
  + 4.4 Preference utilitarianism[[edit](/index.php?title=(none)&action=edit&section=12)]
* 5 More varieties of utilitarianism[[edit](/index.php?title=(none)&action=edit&section=13)]
  + 5.1 Negative utilitarianism[[edit](/index.php?title=(none)&action=edit&section=14)]
  + 5.2 Motive utilitarianism[[edit](/index.php?title=(none)&action=edit&section=15)]
* 6 Criticisms[[edit](/index.php?title=(none)&action=edit&section=16)]
  + 6.1 Ignores justice[[edit](/index.php?title=(none)&action=edit&section=17)]
  + 6.2 Calculating utility is self-defeating[[edit](/index.php?title=(none)&action=edit&section=18)]
  + 6.3 Predicting consequences[[edit](/index.php?title=(none)&action=edit&section=19)]
  + 6.4 Too demanding[[edit](/index.php?title=(none)&action=edit&section=20)]
  + 6.5 Aggregating utility[[edit](/index.php?title=(none)&action=edit&section=21)]
  + 6.6 Individual criticisms[[edit](/index.php?title=(none)&action=edit&section=22)]
    - 6.6.1 Karl Marx's criticisms[[edit](/index.php?title=(none)&action=edit&section=23)]
    - 6.6.2 John Taurek's criticism[[edit](/index.php?title=(none)&action=edit&section=24)]
    - 6.6.3 John Paul II's personality criticism[[edit](/index.php?title=(none)&action=edit&section=25)]

## Etymology[[edit](/index.php?title=(none)&action=edit&section=1)]

As to the origin of the word 'Utilitarianism' Mill acknowledged in a footnote that, though "believing himself to be the first person who brought the word utilitarian into use, he did not invent it. Rather, he adopted it from a passing expression in [Galt's](/wiki/John_Galt_(novelist)) [*Annals of the Parish*](/wiki/Annals_of_the_Parish)."[[1]](#cite_note-1) Mill seems to have been unaware that Bentham had previously used the word 'utilitarian' in his 1781 letter to George Wilson and in Bentham's 1802 letter to Dumont he had claimed that 'Utilitarian' was the proper name for his new creed.

## Historical background[[edit](/index.php?title=(none)&action=edit&section=2)]

[Template:See also](/wiki/Template:See_also) The importance of [happiness](/wiki/Happiness) as an end for humans has long been recognized. Forms of [hedonism](/wiki/Hedonism) were put forward by [Aristippus](/wiki/Aristippus) and [Epicurus](/wiki/Epicurus); [Aristotle](/wiki/Aristotle) argued that [eudaimonia](/wiki/Eudaimonia) is the highest human good and [Augustine](/wiki/Augustine_of_Hippo) wrote that "all men agree in desiring the last end, which is happiness." Happiness was also explored in depth by [Aquinas](/wiki/Thomas_Aquinas).[[2]](#cite_note-2)[[3]](#cite_note-3)[[4]](#cite_note-4)[[5]](#cite_note-5)[[6]](#cite_note-6) Different varieties of consequentialism also existed in the ancient and medieval world, like the [state consequentialism](/wiki/State_consequentialism) of [Mohism](/wiki/Mohism) or the political philosophy of [Niccolò Machiavelli](/wiki/Niccolò_Machiavelli). Mohist consequentialism advocated communitarian moral goods including political stability, population growth, and wealth, but did not support the utilitarian notion of maximizing individual happiness.[[7]](#cite_note-7) Machiavelli was also an exponent of consequentialism. He believed that the actions of a state, however cruel or ruthless they may be, must contribute towards the common good of a society.[[8]](#cite_note-8) Utilitarianism as a distinct ethical position only emerged in the eighteenth century.

Although utilitarianism is usually thought to start with [Jeremy Bentham](/wiki/Jeremy_Bentham), there were earlier writers who presented theories that were strikingly similar. In [*An Enquiry Concerning the Principles of Morals*](/wiki/An_Enquiry_Concerning_the_Principles_of_Morals), [David Hume](/wiki/David_Hume) writes:[[9]](#cite_note-9) [Template:Quote](/wiki/Template:Quote)

Hume studied under [Francis Hutcheson](/wiki/Francis_Hutcheson_(philosopher)), and it was he who first introduced a key utilitarian phrase. In *An Inquiry into the Original of Our Ideas of Beauty and Virtue* (1725), Hutcheson says[[10]](#cite_note-10) when choosing the most moral action, virtue is in proportion to the number of people a particular action brings happiness to. In the same way, moral evil, or vice, is proportionate to the number of people made to suffer. The best action is the one that procures the greatest happiness for the greatest numbers—and the worst is the one that causes the most misery.

In the first three editions of the book, Hutcheson included various mathematical algorithms "...to compute the Morality of any Actions." In this, he pre-figured the [hedonic calculus](/wiki/Felicific_calculus) of Bentham.

Some claim that [John Gay](/wiki/John_Gay_(philosopher)) developed the first systematic theory of utilitarian ethics.[[11]](#cite_note-11) In *Concerning the Fundamental Principle of Virtue or Morality* (1731), Gay argues that:[[12]](#cite_note-12)[Template:Quote](/wiki/Template:Quote)

This pursuit of happiness is given a theological basis:[[13]](#cite_note-13)[Template:Quote](/wiki/Template:Quote)

[thumb|*Modern Utilitarianism* by](/wiki/File:Modern_Utiitarianism_by_Birks.png) [Thomas Rawson Birks](/wiki/Thomas_Rawson_Birks) 1874 Gay's theological utilitarianism was developed and popularized by [William Paley](/wiki/William_Paley). It has been claimed that Paley was not a very original thinker and that the philosophical part of his treatise on ethics is "an assemblage of ideas developed by others and is presented to be learned by students rather than debated by colleagues. "[[14]](#cite_note-14) Nevertheless, his book *The Principles of Moral and Political Philosophy* (1785) was a required text at Cambridge[[14]](#cite_note-14) and Smith says that Paley's writings were "once as well known in American colleges as were the readers and spellers of William McGuffey and Noah Webster in the elementary schools."[[15]](#cite_note-15) Although now largely missing from the philosophical canon, Schneewind writes that "utilitarianism first became widely known in England through the work of William Paley."[[16]](#cite_note-16) The now forgotten significance of Paley can be judged from the title of [Thomas Rawson Birks's](/wiki/Thomas_Rawson_Birks) 1874 work *Modern Utilitarianism or the Systems of Paley, Bentham and Mill Examined and Compared*.

Apart from restating that happiness as an end is grounded in the nature of God, Paley also discusses the place of rules. He writes:[[17]](#cite_note-17)[Template:Quote](/wiki/Template:Quote)

## Classical utilitarianism[[edit](/index.php?title=(none)&action=edit&section=3)]

### Jeremy Bentham[[edit](/index.php?title=(none)&action=edit&section=4)]

[Template:Main](/wiki/Template:Main) [thumb|right|](/wiki/File:Jeremy_Bentham_by_Henry_William_Pickersgill_detail.jpg)[Jeremy Bentham](/wiki/Jeremy_Bentham). Bentham's book *An Introduction to the Principles of Morals and Legislation* was printed in 1780 but not published until 1789. It is possible that Bentham was spurred on to publish after he saw the success of Paley's *The Principles of Moral and Political Philosophy*.[[18]](#cite_note-18) Bentham's book was not an immediate success[[19]](#cite_note-19) but his ideas were spread further when [Pierre Étienne Louis Dumont](/wiki/Pierre_Étienne_Louis_Dumont) translated edited selections from a variety of Bentham's manuscripts into French. *Traité de legislation civile et pénale* was published in 1802 and then later retranslated back into English by Hildreth as *The Theory of Legislation*, although by this time significant portions of Dumont's work had already been retranslated and incorporated into [Sir John Bowring's](/wiki/John_Bowring) edition of Bentham's works, which was issued in parts between 1838 and 1843.

Bentham's work opens with a statement of the principle of utility:[[20]](#cite_note-20) [Template:Quote](/wiki/Template:Quote)

In Chapter IV, Bentham introduces a method of calculating the value of pleasures and pains, which has come to be known as the hedonic calculus. Bentham says that the value of a pleasure or pain, considered by itself, can be measured according to its intensity, duration, certainty/uncertainty and propinquity/remoteness. In addition, it is necessary to consider "the tendency of any act by which it is produced" and, therefore, to take account of the act's fecundity, or the chance it has of being followed by sensations of the same kind and its purity, or the chance it has of not being followed by sensations of the opposite kind. Finally, it is necessary to consider the extent, or the number of people affected by the action.

Perhaps aware that Hutcheson eventually removed his algorithms for calculating the greatest happiness because they "appear'd useless, and were disagreeable to some readers",[[21]](#cite_note-21) Bentham contends that there is nothing novel or unwarranted about his method, for "in all this there is nothing but what the practice of mankind, wheresoever they have a clear view of their own interest, is perfectly conformable to."

Rosen warns that descriptions of utilitarianism can bear "little resemblance historically to utilitarians like Bentham and [J. S. Mill](/wiki/John_Stuart_Mill)" and can be more "a crude version of act utilitarianism conceived in the twentieth century as a straw man to be attacked and rejected."[[22]](#cite_note-22) It is a mistake to think that Bentham is not concerned with rules. His seminal work is concerned with the principles of legislation and the hedonic calculus is introduced with the words "Pleasures then, and the avoidance of pains, are the ends that the legislator has in view." In Chapter VII, Bentham says: "The business of government is to promote the happiness of the society, by punishing and rewarding… In proportion as an act tends to disturb that happiness, in proportion as the tendency of it is pernicious, will be the demand it creates for punishment."

The question then arises as to when, if at all, it might legitimate to break the law. This is considered in *The Theory of Legislation*, where Bentham distinguishes between evils of the first and second orders. Those of the first order are the more immediate consequences; those of the second are when the consequences spread through the community causing "alarm" and "danger".

It is true there are cases in which, if we confine ourselves to the effects of the first order, the good will have an incontestable preponderance over the evil. Were the offence considered only under this point of view, it would not be easy to assign any good reasons to justify the rigour of the laws. Every thing depends upon the evil of the second order; it is this which gives to such actions the character of crime, and which makes punishment necessary. Let us take, for example, the physical desire of satisfying hunger. Let a beggar, pressed by hunger, steal from a rich man's house a loaf, which perhaps saves him from starving, can it be possible to compare the good which the thief acquires for himself, with the evil which the rich man suffers? … It is not on account of the evil of the first order that it is necessary to erect these actions into offences, but on account of the evil of the second order.[[23]](#cite_note-23)

### John Stuart Mill[[edit](/index.php?title=(none)&action=edit&section=5)]

[Template:Main](/wiki/Template:Main) Mill was brought up as a Benthamite with the explicit intention that he would carry on the cause of utilitarianism.[[24]](#cite_note-24) Mill's book [*Utilitarianism*](/wiki/Utilitarianism_(book)) first appeared as a series of three articles published in [*Fraser's Magazine*](/wiki/Fraser's_Magazine) in 1861 and was reprinted as a single book in 1863.[Template:Citation needed](/wiki/Template:Citation_needed)

#### Higher and lower pleasures[[edit](/index.php?title=(none)&action=edit&section=6)]

Mill rejects a purely quantitative measurement of utility and says:[[25]](#cite_note-25)[Template:Quote](/wiki/Template:Quote)

Mill notes that, contrary to what its critics might say, there is "no known Epicurean theory of life which does not assign to the pleasures of the intellect… a much higher value as pleasures than to those of mere sensation." However, he accepts that this is usually because the intellectual pleasures are thought to have circumstantial advantages, i.e. "greater permanency, safety, uncostliness, &c." Instead, Mill will argue that some pleasures are intrinsically better than others.

The accusation that hedonism is "doctrine worthy only of swine" has a long history. In [Nicomachean Ethics](/wiki/Nicomachean_Ethics) (Book 1 Chapter 5), [Aristotle](/wiki/Aristotle) says that identifying the good with pleasure is to prefer a life suitable for beasts. The theological utilitarians had the option of grounding their pursuit of happiness in the will of God; the hedonistic utilitarians needed a different defense. Mill's approach is to argue that the pleasures of the intellect are intrinsically superior to physical pleasures.

Few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance of a beast's pleasures; no intelligent human being would consent to be a fool, no instructed person would be an ignoramus, no person of feeling and conscience would be selfish and base, even though they should be persuaded that the fool, the dunce, or the rascal is better satisfied with his lot than they are with theirs… A being of higher faculties requires more to make him happy, is capable probably of more acute suffering, an certainly accessible to it at more points, than one of an inferior type; but in spite of these liabilities, he can never really wish to sink into what he feels to be a lower grade of existence… It is better to be a human being dissatisfied than a pig satisfied; better to be [Socrates](/wiki/Socrates) dissatisfied than a fool satisfied. And if the fool, or the pig, are of a different opinion, it is because they only know their own side of the question…[[26]](#cite_note-26)

Mill argues that if people who are "competently acquainted" with two pleasures show a decided preference for one even if it be accompanied by more discontent and "would not resign it for any quantity of the other", then it is legitimate to regard that pleasure as being superior in quality. Mill recognises that these "competent judges" will not always agree, and states that, in cases of disagreement, the judgment of the majority is to be accepted as final. Mill also acknowledges that "many who are capable of the higher pleasures, occasionally, under the influence of temptation, postpone them to the lower. But this is quite compatible with a full appreciation of the intrinsic superiority of the higher." Mill says that this appeal to those who have experienced the relevant pleasures is no different from what must happen when assessing the quantity of pleasure, for there is no other way of measuring "the acutest of two pains, or the intensest of two pleasurable sensations." "It is indisputable that the being whose capacities of enjoyment are low, has the greatest chance of having them fully satisfied; and a highly-endowed being will always feel that any happiness which he can look for, as the world is constitute, is imperfect."[[27]](#cite_note-27)

#### Mill's "proof" of the principle of utility[[edit](/index.php?title=(none)&action=edit&section=7)]

In Chapter Four of *Utilitarianism*, Mill considers what proof can be given for the principle of utility. He says:[[28]](#cite_note-28)[Template:Quote](/wiki/Template:Quote)

It is usual[[29]](#cite_note-29) to say that Mill is committing a number of fallacies. He is accused of committing the [naturalistic fallacy](/wiki/Naturalistic_fallacy), because he is trying to deduce what people ought to do from what they do in fact do; the [fallacy of equivocation](/wiki/Fallacy_of_equivocation), because he moves from the fact that (1) something is desirable, i.e. is capable of being desired, to the claim that (2) it is desirable, i.e. that it ought to be desired; and the [fallacy of composition](/wiki/Fallacy_of_composition), because the fact that people desire their own happiness does not imply that the aggregate of all persons will desire the general happiness.

Such allegations began to emerge in Mill’s lifetime, shortly after the publication of *Utilitarianism*, and persisted for well over a century, though the tide has been turning in recent discussions.

A defense of Mill against all three charges, with a chapter devoted to each, can be found in Necip Fikri Alican’s *Mill’s Principle of Utility: A Defense of John Stuart Mill’s Notorious Proof* (1994). This is the first, and remains the only, book-length treatment of the subject matter. Yet the alleged fallacies in the proof continue to attract scholarly attention in journal articles and book chapters.

Hall[[30]](#cite_note-30) and Popkin[[31]](#cite_note-31) defend Mill against this accusation pointing out that he begins Chapter Four by asserting that "questions of ultimate ends do not admit of proof, in the ordinary acceptation of the term" and that this is "common to all first principles." According to Hall and Popkin, therefore, Mill does not attempt to "establish that what people do desire is desirable but merely attempts to make the principles acceptable."[[29]](#cite_note-29) The type of "proof" Mill is offering "consists only of some considerations which, Mill thought, might induce an honest and reasonable man to accept utilitarianism."[[29]](#cite_note-29) Having claimed that people do, in fact, desire happiness, Mill now has to show that it is the *only* thing they desire. Mill anticipates the objection that people desire other things such as virtue. He argues that whilst people might start desiring virtue as a *means* to happiness, eventually, it becomes part of someone's happiness and is then desired as an end in itself.

[Template:Quote](/wiki/Template:Quote)

[Template:Quote](/wiki/Template:Quote)

## Twentieth-century developments[[edit](/index.php?title=(none)&action=edit&section=8)]

### Ideal utilitarianism[[edit](/index.php?title=(none)&action=edit&section=9)]

The description ideal utilitarianism was first used by [Hastings Rashdall](/wiki/Hastings_Rashdall) in [*The Theory of Good and Evil*](/wiki/The_Theory_of_Good_and_Evil) (1907), but it is more often associated with [G. E. Moore](/wiki/G._E._Moore). In *Ethics* (1912), Moore rejected a purely hedonistic utilitarianism and argued that there is a range of values that might be maximized. Moore's strategy was to show that it is intuitively implausible that pleasure is the sole measure of what is good. He says that such an assumption:[[32]](#cite_note-32)[Template:Quote](/wiki/Template:Quote)

Moore admits that it is impossible to prove the case either way, but he believed that it was intuitively obvious that even if the amount of pleasure stayed the same a world that contained such things as beauty and love would be a better world. He adds that, if a person was to take the contrary view, then "I think it is self-evident that he would be wrong."[[32]](#cite_note-32)

### Act and rule utilitarianism[[edit](/index.php?title=(none)&action=edit&section=10)]

In the mid-twentieth century a number of philosophers focused on the place of rules in utilitarian thinking.[[33]](#cite_note-33) It was already accepted that it is necessary to use rules to help you choose the right action because the problems of calculating the consequences on each and every occasion would almost certainly result in you frequently choosing something less than the best course of action. Paley had justified the use of rules and Mill says:[[34]](#cite_note-34) [Template:Quote](/wiki/Template:Quote)

However, rule utilitarianism proposes a more central role for rules that was thought to rescue the theory from some of its more devastating criticisms, particularly problems to do with justice and promise keeping. Throughout the 1950s and 1960s, articles were published both for and against the new form of utilitarianism, and through this debate the theory we now call rule utilitarianism was created. In an introduction to an anthology of these articles, the editor was able to say: "The development of this theory was a dialectical process of formulation, criticism, reply and reformulation; the record of this process well illustrates the co-operative development of a philosophical theory."[[35]](#cite_note-35) Smart[[36]](#cite_note-36) and McCloskey<ref name=McCloskey1957>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> initially used the terms 'extreme' and 'restricted' utilitarianism but eventually everyone settled on the terms 'act' and 'rule' utilitarianism.

The essential difference is in what determines whether or not an action is the right action. [Act utilitarianism](/wiki/Act_utilitarianism) maintains that an action is right if it maximises utility; rule utilitarianism maintains that an action is right if it conforms to a rule that maximises utility.

In 1956, Urmson published an influential article[[37]](#cite_note-37) arguing that Mill justified rules on utilitarian principles. From then on, articles have debated this interpretation of Mill. In all probability, it was not a distinction that Mill was particularly trying to make and so the evidence in his writing is inevitably mixed. A collection of Mill's writing published in 1977 includes a letter in which he says:[[38]](#cite_note-38)[Template:Quote](/wiki/Template:Quote)

This seems to tip the balance in favour of saying that Mill is best classified as an act utilitarian.

Some school level textbooks and at least one UK examination board[[39]](#cite_note-39) make a further distinction between strong and weak rule utilitarianism. However, it is not clear that this distinction is made in the academic literature.

It has been argued that rule utilitarianism collapses into act utilitarianism, because for any given rule, in the case where breaking the rule produces more utility, the rule can be refined by the addition of a sub-rule that handles cases like the exception.[[40]](#cite_note-40) This process holds for all cases of exceptions, and so the "rules" have as many "sub-rules" as there are exceptional cases, which, in the end, makes an agent seek out whatever outcome produces the maximum utility.[[41]](#cite_note-41)

### Two-level utilitarianism[[edit](/index.php?title=(none)&action=edit&section=11)]

[Template:Main](/wiki/Template:Main) In *Principles* (1973),[[42]](#cite_note-42) [R. M. Hare](/wiki/R._M._Hare) accepts that rule utilitarianism collapses into act utilitarianism but claims that this is a result of allowing the rules to be "as specific and un-general as we please." He argues that one of the main reasons for introducing rule utilitarianism was to do justice to the general rules that people need for moral education and character development and he proposes that "a difference between act-utilitarianism and rule-utilitarianism can be introduced by limiting the specificity of the rules, i.e., by increasing their generality."[[42]](#cite_note-42)[Template:Rp](/wiki/Template:Rp) This distinction between a "specific rule utilitarianism" (which collapses into act utilitarianism) and "general rule utilitarianism" forms the basis of Hare's two-level utilitarianism.

When we are "playing God or the ideal observer", we use the specific form, and we will need to do this when we are deciding what general principles to teach and follow. When we are "inculcating" or in situations where the biases of our human nature are likely to prevent us doing the calculations properly, then we should use the more general rule utilitarianism.

Hare argues that in practice, most of the time, we should be following the general principles:[[42]](#cite_note-42)[Template:Rp](/wiki/Template:Rp)

[Template:Quote](/wiki/Template:Quote)

In *Moral Thinking* (1981), Hare illustrated the two extremes. The "archangel" is the hypothetical person who has perfect knowledge of the situation and no personal biases or weaknesses and always uses critical moral thinking to decide the right thing to do; the "prole" is the hypothetical person who is completely incapable of critical thinking and uses nothing but intuitive moral thinking and, of necessity, has to follow the general moral rules they have been taught or learned through imitation.[[43]](#cite_note-43) It is not that some people are archangels and others proles, but rather that "we all share the characteristics of both to limited and varying degrees and at different times."[[43]](#cite_note-43) Hare does not specify when we should think more like an "archangel" and more like a "prole" as this will, in any case, vary from person to person. However, the critical moral thinking underpins and informs the more intuitive moral thinking. It is responsible for formulating and, if necessary, reformulating the general moral rules. We also switch to critical thinking when trying to deal with unusual situations or in cases where the intuitive moral rules give conflicting advice.

### Preference utilitarianism[[edit](/index.php?title=(none)&action=edit&section=12)]

Preference utilitarianism was first put forward in 1977 by [John Harsanyi](/wiki/John_Harsanyi) in *Morality and the theory of rational behaviour*,[[44]](#cite_note-44) but preference utilitarianism is more commonly associated with [R. M. Hare](/wiki/R._M._Hare),[[43]](#cite_note-43) [Peter Singer](/wiki/Peter_Singer)[[45]](#cite_note-45) and [Richard Brandt](/wiki/Richard_Brandt).[[46]](#cite_note-46) Harsanyi claimed that his theory is indebted to [Adam Smith](/wiki/Adam_Smith), who equated the moral point of view with that of an impartial but sympathetic observer; to [Kant](/wiki/Immanuel_Kant), who insisted on the criterion of universality, which may also be described as a criterion of reciprocity; to the classical utilitarians who made maximising social utility the basic criterion of morality; and to "the modern theory of rational behaviour under risk and uncertainty, usually described as Bayesian [decision theory](/wiki/Decision_theory)".[[44]](#cite_note-44)[Template:Rp](/wiki/Template:Rp)

Harsanyi rejects hedonistic utilitarianism as being dependent on an outdated psychology saying that it is far from obvious that everything we do is motivated by a desire to maximise pleasure and minimise pain. He also rejects ideal utilitarianism because "it is certainly not true as an empirical observation that people's only purpose in life is to have 'mental states of intrinsic worth'."[[44]](#cite_note-44)[Template:Rp](/wiki/Template:Rp)

According to Harsanyi, "preference utilitarianism is the only form of utilitarianism consistent with the important philosophical principle of preference autonomy. By this I mean the principle that, in deciding what is good and what is bad for a given individual, the ultimate criterion can only be his own wants and his own preferences."[[44]](#cite_note-44)[Template:Rp](/wiki/Template:Rp)

Harsanyi adds two caveats. People sometimes have irrational preferences. To deal with this, Harsanyi distinguishes between "manifest" preferences and "true" preferences. The former are those "manifested by his observed behaviour, including preferences possibly based on erroneous factual beliefs, or on careless logical analysis, or on strong emotions that at the moment greatly hinder rational choice" whereas the latter are "the preferences he would have if he had all the relevant factual information, always reasoned with the greatest possible care, and were in a state of mind most conducive to rational choice."[[44]](#cite_note-44)[Template:Rp](/wiki/Template:Rp) It is the latter that preference utilitarianism tries to satisfy.

The second caveat is that antisocial preferences, such as sadism, envy and resentment, have to be excluded. Harsanyi achieves this by claiming that such preferences partially exclude those people from the moral community: [Template:Quote](/wiki/Template:Quote)

## More varieties of utilitarianism[[edit](/index.php?title=(none)&action=edit&section=13)]

### Negative utilitarianism[[edit](/index.php?title=(none)&action=edit&section=14)]

[Template:Main](/wiki/Template:Main) In *The Open Society and its Enemies* (1945), [Karl Popper](/wiki/Karl_Popper) argued that the principle "maximize pleasure" should be replaced by "minimize pain". He thought "it is not only impossible but very dangerous to attempt to maximize the pleasure or the happiness of the people, since such an attempt must lead to totalitarianism."[[47]](#cite_note-47) He claimed that:[[48]](#cite_note-48) [Template:Quote](/wiki/Template:Quote)

The actual term *negative utilitarianism* was introduced by [R.N.Smart](/wiki/Ninian_Smart) as the title to his 1958 reply to Popper<ref name=Smart1958>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> in which he argued that the principle would entail seeking the quickest and least painful method of killing the entirety of humanity.

Negative *total* utilitarianism, in contrast, tolerates suffering that can be compensated within the same person.[[49]](#cite_note-49)[[50]](#cite_note-50) Negative *preference* utilitarianism avoids the problem of moral killing with reference to existing preferences that such killing would violate, while it still demands a justification for the creation of new lives.[[51]](#cite_note-51) A possible justification is the reduction of the average level of preference-frustration.[[52]](#cite_note-52) Others see negative utilitarianism as a branch within classical utilitarianism, which assigns a higher weight to the avoidance of suffering than to the promotion of happiness.[[53]](#cite_note-53) The moral weight of suffering can be increased by using a "compassionate" utilitarian metric, so that the result is the same as in [prioritarianism](/wiki/Prioritarianism).[[54]](#cite_note-54) Pessimistic representatives of negative utilitarianism can be found in the environment of [Buddhism](/wiki/Buddhism).[[55]](#cite_note-55)

### Motive utilitarianism[[edit](/index.php?title=(none)&action=edit&section=15)]

Motive utilitarianism was first proposed by [Robert Merrihew Adams](/wiki/Robert_Merrihew_Adams) in 1976.[[56]](#cite_note-56) Whereas act utilitarianism requires us to choose our actions by calculating which action will maximize [utility](/wiki/Utility) and rule utilitarianism requires us to implement rules which will, on the whole, maximize utility, motive utilitarianism "has the utility calculus being used to select motives and dispositions according to their general felicific effects, and those motives and dispositions then dictate our choices of actions."[[57]](#cite_note-57) The arguments for moving to some form of motive utilitarianism at the personal level can be seen as mirroring the arguments for moving to some form of rule utilitarianism at the social level.[[58]](#cite_note-58) Adams refers to Sidgwick's observation that "Happiness (general as well as individual) is likely to be better attained if the extent to which we set ourselves consciously to aim at it be carefully restricted."[[59]](#cite_note-59) Trying to apply the utility calculation on each and every occasion is likely to lead to a sub-optimal outcome. Applying carefully selected rules at the social level and encouraging appropriate motives at the personal level is, so it is argued, likely to lead to a better overall outcome even if on some individual occasions it leads to the wrong action when assessed according to act utilitarian standards.

Adams illustrates his theory by telling a fictitious story about Jack, a lover of art, visiting Chartres cathedral. Jack is motivated to see, as nearly as possible, everything in the cathedral. However, there were some things in the cathedral that, on their own, didn't interest him much. On act utilitarian grounds he should have ignored them. Spending more time in the cathedral than he had originally planned resulted in him missing his dinner, doing several hours of night driving, which he hates, and having trouble finding a place to sleep. Adams argues that Jack will only have skipped the less interesting bits of the cathedral if "he had been less interested in seeing everything in the cathedral than in maximizing utility. And it is plausible to suppose that if his motivation had been different in that respect, he would have enjoyed the cathedral much less."[[60]](#cite_note-60) Adams concludes that "right action, by act-utilitarian standards, and right motivation, by motive-utilitarian standards, are incompatible in some cases."[[61]](#cite_note-61) The necessity of this conclusion is rejected by [Fred Feldman](/wiki/Fred_Feldman_(philosopher)) who argues that "the conflict in question results from an inadequate formulation of the utilitarian doctrines; motives play no essential role in it…(and that)… Precisely the same sort of conflict arises even when MU is left out of consideration and AU is applied by itself."[[62]](#cite_note-62) Instead, [Feldman](/wiki/Fred_Feldman_(philosopher)) proposes a variant of act utilitarianism that results in there being no conflict between it and motive utilitarianism.

## Criticisms[[edit](/index.php?title=(none)&action=edit&section=16)]

Because utilitarianism is not a single theory but a cluster of related theories that have developed over two hundred years, criticisms can be made for different reasons and have different targets.

### Ignores justice[[edit](/index.php?title=(none)&action=edit&section=17)]

As Rosen[[18]](#cite_note-18) has pointed out, claiming that act utilitarians are not concerned about having rules is to set up a "straw man". Similarly, [Hare](/wiki/R._M._Hare) refers to "the crude caricature of act utilitarianism which is the only version of it that many philosophers seem to be acquainted with."[[63]](#cite_note-63) Given what Bentham says about second order evils[[64]](#cite_note-64) it would be a serious misrepresentation to say that he and similar act utilitarians would be prepared to punish an innocent person for the greater good. Nevertheless, whether they would agree or not, this is what critics of utilitarianism claim is entailed by the theory. A classic version of this criticism was given by H. J. McCloskey:<ref name=McCloskey1957/> [Template:Quote](/wiki/Template:Quote)

By "extreme" utilitarian, McCloskey is referring to what later came to be called "act" utilitarianism. Whilst this story might be quoted as part of a justification for moving from act to rule utilitarianism McCloskey anticipates this and points out that each rule has to be judged on its utility and it is not at all obvious that a rule with exceptions has less utility. The above story invites the reply that the sheriff would not frame the innocent because of the rule "do not punish an innocent person"; it also invites the reply that these issues need to be resolved, and riots might very well have positive utility in the long run by drawing attention and thus resources to the racial situation. However, McCloskey asks, what about the rule "punish an innocent person when and only when to do so is not to weaken the existing institution of punishment and when the consequences of doing so are valuable"?

In a later article, McCloskey says:[[65]](#cite_note-65)[Template:Quote](/wiki/Template:Quote)

### Calculating utility is self-defeating[[edit](/index.php?title=(none)&action=edit&section=18)]

An early criticism, which was addressed by Mill, is that if time is taken to calculate the best course of action it is likely that the opportunity to take the best course of action will already have passed. Mill responded that there had been ample time to calculate the likely effects:[[66]](#cite_note-66)[Template:Quote](/wiki/Template:Quote)

More recently, Hardin has made the same point. "It should embarrass philosophers that they have ever taken this objection seriously. Parallel considerations in other realms are dismissed with eminently good sense. Lord Devlin notes, 'if the reasonable man "[worked to rule](/wiki/Work-to-rule)" by perusing to the point of comprehension every form he was handed, the commercial and administrative life of the country would creep to a standstill.[Template:'"](/wiki/Template:'%22)[[67]](#cite_note-67) It is such considerations that lead even act utilitarians to rely on "rules of thumb", as [Smart](/wiki/J._J._C._Smart)[[68]](#cite_note-68) has called them. The objection arises when utilitarianism is mistakenly taken to be a decision-making procedure rather than a criterion of what is right.

### Predicting consequences[[edit](/index.php?title=(none)&action=edit&section=19)]

Some argue that it is impossible to do the calculation that utilitarianism requires because consequences are inherently unknowable. [Daniel Dennett](/wiki/Daniel_Dennett) describes this as the [Three Mile Island](/wiki/Three_Mile_Island_accident) effect.[[69]](#cite_note-69) [Dennett](/wiki/Daniel_Dennett) points out that not only is it impossible to assign a precise utility value to the incident, it is impossible to know whether, ultimately, the near-meltdown that occurred was a good or bad thing. He suggests that it would have been a good thing if plant operators learned lessons that prevented future serious incidents.

Russell Hardin rejects such arguments. He argues that it is possible to distinguish the moral impulse of utilitarianism (which is "to define the right as good consequences and to motivate people to achieve these") from our ability to correctly apply rational principles which will among other things "depend on the perceived facts of the case and on the particular moral actor's mental equipment."[[70]](#cite_note-70) The fact that the latter is limited and can change doesn't mean that the former has to be rejected. "If we develop a better system for determining relevant causal relations so that we are able to choose actions that better produce our intended ends, it does not follow that we then must change our ethics. The moral impulse of utilitarianism is constant, but our decisions under it are contingent on our knowledge and scientific understanding."[[67]](#cite_note-67) From the beginning, utilitarianism has recognized that certainty in such matters is unobtainable and both Bentham and Mill said that it was necessary to rely on the *tendencies* of actions to bring about consequences. [G. E. Moore](/wiki/G._E._Moore) writing in 1903 said:[[71]](#cite_note-71)[Template:Quote](/wiki/Template:Quote)

### Too demanding[[edit](/index.php?title=(none)&action=edit&section=20)]

Act utilitarianism not only requires everyone to do what they can to maximise utility, but to do so without any favouritism. Mill says, "As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator."[[66]](#cite_note-66) Critics say that this combination of requirements leads to utilitarianism making unreasonable demands. The well-being of strangers counts just as much as that of friends, family or self. "What makes this requirement so demanding is the gargantuan number of strangers in great need of help and the indefinitely many opportunities to make sacrifices to help them."[[72]](#cite_note-72) As Shelly Kagan says, "Given the parameters of the actual world, there is no question that …(maximally)… promoting the good would require a life of hardship, self-denial, and austerity…a life spent promoting the good would be a severe one indeed."[[73]](#cite_note-73) Hooker describes two aspects to the problem: act utilitarianism requires *huge* sacrifices from those who are relatively better off and also requires sacrifice of your own good even when the aggregate good will be only *slightly* increased.[[74]](#cite_note-74) Another way of highlighting the complaint is to say that in utilitarianism, "there is no such thing as morally permissible self-sacrifice that goes above and beyond the call of duty."[[74]](#cite_note-74) Mill was quite clear about this, "A sacrifice which does not increase, or tend to increase, the sum total of happiness, it considers as wasted."[[66]](#cite_note-66) One response to the problem is to accept its demands. This is the view taken by Peter Singer, who says: "No doubt we do instinctively prefer to help those who are close to us. Few could stand by and watch a child drown; many can ignore the avoidable deaths of children in Africa or India. The question, however, is not what we usually do, but what we ought to do, and it is difficult to see any sound moral justification for the view that distance, or community membership, makes a crucial difference to our obligations."[[75]](#cite_note-75) Others argue that a moral theory that is so contrary to our deeply held moral convictions must either be rejected or modified.[[76]](#cite_note-76) There have been various attempts to modify utilitarianism to escape its seemingly over-demanding requirements.[[77]](#cite_note-77) One approach is to drop the demand that utility be maximized. In [Satisficing Consequentialism,](/wiki/Satisficing) Michael Slote argues for a form of utilitarianism where "an act might qualify as morally right through having good enough consequences, even though better consequences could have been produced."[[78]](#cite_note-78) One advantage of such a system is that it would be able to accommodate the notion of supererogatory actions.

Samuel Scheffler takes a different approach and amends the requirement that everyone be treated the same.[[79]](#cite_note-79) In particular, Scheffler suggests that there is an "agent-centered prerogative" such that when the overall utility is being calculated it is permitted to count our own interests more heavily than the interests of others. Kagan suggests that such a procedure might be justified on the grounds that "a general requirement to promote the good would lack the motivational underpinning necessary for genuine moral requirements" and, secondly, that personal independence is necessary for the existence of commitments and close personal relations and that "the value of such commitments yields a positive reason for preserving within moral theory at least some moral independence for the personal point of view."[[80]](#cite_note-80) Robert Goodin takes yet another approach and argues that the demandingness objection can be "blunted" by treating utilitarianism as a guide to public policy rather than one of individual morality. He suggests that many of the problems arise under the traditional formulation because the conscientious utilitarian ends up having to make up for the failings of others and so contributing more than their fair share.[[81]](#cite_note-81) Harsanyi argues that the objection overlooks the fact that "people attach considerable utility to freedom from unduly burdensome moral obligations… most people will prefer a society with a more relaxed moral code, and will feel that such a society will achieve a higher level of average utility—even if adoption of such a moral code should lead to some losses in economic and cultural accomplishments (so long as these losses remain within tolerable limits). This means that utilitarianism, if correctly interpreted, will yield a moral code with a standard of acceptable conduct very much below the level of highest moral perfection, leaving plenty of scope for supererogatory actions exceeding this minimum standard."[[82]](#cite_note-82)

### Aggregating utility[[edit](/index.php?title=(none)&action=edit&section=21)]

The objection that "utilitarianism does not take seriously the distinction between persons"[[83]](#cite_note-83) came to prominence in 1971 with the publication of [John Rawls'](/wiki/John_Rawls) [*A Theory of Justice*](/wiki/A_Theory_of_Justice). The concept is also important in [animal rights](/wiki/Animal_rights) advocate [Richard Ryder's](/wiki/Richard_D._Ryder) rejection of utilitarianism, in which he talks of the "boundary of the individual", through which neither pain nor pleasure may pass.[[84]](#cite_note-84) However, a similar objection was noted in 1970 by [Thomas Nagel](/wiki/Thomas_Nagel) (who claimed that consequentialism "treats the desires, needs, satisfactions, and dissatisfactions of distinct persons as if they were the desires, etc., of a mass person"[[85]](#cite_note-85)), and even earlier by David Gauthier, who wrote that utilitarianism supposes "that mankind is a super-person, whose greatest satisfaction is the objective of moral action. . . . But this is absurd. Individuals have wants, not mankind; individuals seek satisfaction, not mankind. A person's satisfaction is not part of any greater satisfaction."[[86]](#cite_note-86) Thus, the aggregation of utility becomes futile as both pain and happiness are intrinsic to and inseparable from the consciousness in which they are felt, rendering impossible the task of adding up the various pleasures of multiple individuals.

A response to this criticism is to point out that whilst seeming to resolve some problems it introduces others. Intuitively, there are many cases where people do want to take the numbers involved into account. As Alastair Norcross has said, "suppose that Homer is faced with the painful choice between saving Barney from a burning building or saving both Moe and Apu from the building…it is clearly better for Homer to save the larger number, precisely because it is a larger number… Can anyone who really considers the matter seriously honestly claim to believe that it is worse that one person die than that the entire sentient population of the universe be severely mutilated? Clearly not."[[87]](#cite_note-87) It may be possible to uphold the distinction between persons whilst still aggregating utility, if it accepted that people can be influenced by [empathy](/wiki/Empathy).[[88]](#cite_note-88) This position is advocated by [Iain King](/wiki/Iain_King),[[89]](#cite_note-89) who has suggested the evolutionary basis of empathy means humans can take into account the interests of other individuals, but only on a one-to-one basis, "since we can only imagine ourselves in the mind of one other person at a time."[[90]](#cite_note-90) King uses this insight to [adapt utilitarianism](/wiki/How_to_Make_Good_Decisions_and_Be_Right_All_the_Time), and it may help reconcile Bentham's philosophy with deontology and virtue ethics.[[91]](#cite_note-91)

### Individual criticisms[[edit](/index.php?title=(none)&action=edit&section=22)]

#### Karl Marx's criticisms[[edit](/index.php?title=(none)&action=edit&section=23)]

[Karl Marx](/wiki/Karl_Marx), in [*Das Kapital*](/wiki/Das_Kapital), writes:[[92]](#cite_note-92)

Not even excepting our philosopher, [Christian Wolff](/wiki/Christian_Wolff_(philosopher)), in no time and in no country has the most homespun commonplace ever strutted about in so self-satisfied a way. The principle of utility was no discovery of Bentham. He simply reproduced in his dull way what [Helvétius](/wiki/Claude_Adrienne_Helvetius) and other Frenchmen had said with esprit in the 18th century. To know what is useful for a dog, one must study dog-nature. This nature itself is not to be deduced from the principle of utility. Applying this to man, he who would criticize all human acts, movements, relations, etc., by the principle of utility, must first deal with human nature in general, and then with human nature as modified in each historical epoch. Bentham makes short work of it. With the driest naivete he takes the modern shopkeeper, especially the English shopkeeper, as the normal man. Whatever is useful to this queer normal man, and to his world, is absolutely useful. This yard-measure, then, he applies to past, present, and future. The Christian religion, e.g., is "useful," "because it forbids in the name of religion the same faults that the penal code condemns in the name of the law." Artistic criticism is "harmful," because it disturbs worthy people in their enjoyment of [Martin Tupper](/wiki/Martin_Tupper), etc. With such rubbish has the brave fellow, with his motto, "null a dies sine linea", piled up mountains of books.

Marx's accusation is twofold. In the first place, he says that the theory of utility is true by definition and thus does not really add anything meaningful. For Marx, a productive inquiry had to investigate what sorts of things are good for people—that is, what our nature, alienated under capitalism, really is. Second, he says that Bentham fails to take account of the *changing* character of people, and hence the changing character of what is good for them. This criticism is especially important for Marx, because he believed that all important statements were contingent upon particular historical conditions.

Marx argues that [human nature](/wiki/Human_nature) is dynamic, so the concept of a single utility for all humans is one-dimensional and not useful. When he decries Bentham's application of the "yard measure" of now to "the past, present and future", he decries the idea that society, and people, have always been, and will always be, as they are now; that is, [essentialism](/wiki/Essentialism). As he sees it, this implication is conservatively used to reinforce institutions he regarded as [reactionary](/wiki/Reactionary). Just because in this moment religion has some positive consequences, says Marx, does not mean that viewed historically it is not a regressive institution that should be abolished.

#### John Taurek's criticism[[edit](/index.php?title=(none)&action=edit&section=24)]

John Taurek has argued that the idea of adding happiness or pleasures across persons is quite unintelligible and that the numbers of persons involved in a situation are morally irrelevant.[[93]](#cite_note-93) Taurek asks whether "we should, in [certain] trade-off situations, consider the relative numbers of people involved as something in itself of significance in determining our course of action[?]" Taurek tells us that "The conclusion I reach is that we should not." Taurek's argument looks at a trade off situation: "The situation is that I have a supply of some life-saving drug. Six people will all certainly die if they are not treated with the drug. But one of the six requires all of the drug if he is to survive. Each of the other five requires only one-fifth of the drug. What ought I to do?" Taurek's basic concern comes down to there being no way to explain what the meaning is of saying that things would be five times worse if the five died than if the one died. "I cannot give a satisfactory account of the meaning of judgments of this kind," he writes (p. 304). He argues that the six persons in this situation, if considered equal in all other respects, should all be given an equal chance of surviving: "I am inclined to treat each person equally by giving each an equal chance to survive." (P. 306.) Each person in the situation can only lose one person's happiness or pleasures. There isn't five times more loss of happiness or pleasure when five die: who would be feeling this happiness or pleasure? "Each person's potential loss has the same significance to me, only as a loss to that person alone. because, by hypothesis, I have an equal concern for each person involved, I am moved to give each of them an equal chance to be spared his loss." (P. 307.) The basic concern here is cogent: while one can understand why more pain or sadness is worse for an individual subject since someone experiences that greater pain or sorrow. But in virtue of what should we take five people's pain or sorrow (all else being equal) as worse if no single person experiences that pain or sorrow? [Parfit](/wiki/Derek_Parfit)[[94]](#cite_note-94) and others[[95]](#cite_note-95) have criticized Taurek's line, and it continues to be discussed.[[96]](#cite_note-96)

#### John Paul II's personality criticism[[edit](/index.php?title=(none)&action=edit&section=25)]

[Pope John Paul II](/wiki/Pope_John_Paul_II), following his personalist philosophy, considered that a danger of utilitarianism is that it tends to make persons, just as much as things, the object of use. "Utilitarianism is a civilization of production and of use, a civilization of things and not of persons, a civilization in which persons are used in the same way as things are used."[[97]](#cite_note-97)