

Utility

Utility in philosophy means something close to pleasure or happiness. According to the philosophy of utilitarianism one ought to act to maximize the total, or in some versions the average,¹ of human utility.

In economics, utility describes not what makes someone happy but what he chooses to do. "The utility to me of an apple is greater than that of an orange" means that, given the choice, I will choose the apple. We expect that people will usually choose what makes them happy, hence expect a close correlation between utility in the economist's sense and utility in the philosopher's sense, but not an identity.

One of the things economists do, when they are not making a point of being objective, value-free scientists, is to draw conclusions about what people ought to do, for instance that they ought to abolish tariffs and price controls.² Those conclusions usually depend on the assumption, stated or unstated, that maximizing utility in the economist's sense will also maximize it in the philosopher's sense. That was clearer a little over a century ago when the economic arguments were being made by an economist, Alfred Marshall, who was not afraid to make explicit the utilitarian foundations of his economic conclusions.

The two senses of utility are correlated but not the same. Suppose you are going to die six months from now. Is your utility greater if you have several months advanced warning, as cancer patients often do, or if your death comes as a complete surprise? Spending several months knowing that you are about to die would be, for most of us, a very unpleasant experience, so if utility is another word for happiness, imagined as a characteristic of what is going on inside your head, the second alternative is almost certainly preferable to the first.

Happiness is not all that matters to people. If one could choose in advance, perhaps by instructions to your doctor, many of us would prefer to know. We have things we would like to get done before dying, things to be said to children, wife, friends, projects to be completed whose completion matters, if only to our sense of having lived a life worth living, arrangements to be made for the future of those dear to us. A close friend spent a good deal of his last few months reducing to something more like order his crowded and cluttered house for the benefit of his wife and daughters.

For another example of tension between what makes you happy and what you choose, consider Robert Nozick's experience machine.³

Someone invents an experience machine; get into it and you will have a fully convincing illusion of experience. The inventor, who somehow knows what your life is going to be like, makes you the following offer:

¹ Utility and Utilitarianism were invented by Jeremy Bentham. Henry Sidgwick seems to be the first person to have raised the question of whether it was the total or average that should be maximized, relevant if one is comparing alternatives with different numbers of people. I offer a partial solution to that problem in "[What Does Optimum Population Mean?](#)" Research in Population Economics, Vol. III (1981), Eds. Simon and Lindert.

² I discuss the issue in [Chapter 15](#) of *Price Theory*.

³ Nozick describes the experience machine in *Anarchy, State and Utopia*. The version of the idea presented here is based on his but modified for my purposes.

Get into my experience machine, spend the rest of your life there, and I will give you the illusion of a life slightly better than the one you would otherwise live. Your average income in the illusion will be a few thousand dollars higher than it would have been in reality, your wife a little prettier, your children slightly better behaved, your promotions a little prompter. Your illusory summers won't be quite as hot or winters quite as cold. Once you are in the machine you will not know that it is an illusion.

Assume you believe his offer. Do you accept it? If not, why not?

If all that matters to you about the world is how it impinges on you, what effect it has on your sensations, you should accept the offer. But for me and, I suspect, many other people, that is not all that matters. I do not merely want the illusion of having written an interesting, enjoyable and original book, I want to have actually done it. I do not just want to think people read my books and are affected by the ideas in them, I want them to actually read and be affected. I don't want just the illusion of wonderful children, I want my wonderful children to actually exist.

I wouldn't touch that machine with a ten-foot pole.

The question is relevant to things more realistic than Nozick's hypothetical. Consider recreational drugs. A lot of us have a gut level feeling that the pleasure from being high on a drug, however intense, is somehow less valid, less real, than the pleasure from accomplishing something, even if it is only winning a game of tennis or climbing a mountain. Feeling good about yourself because you are drunk is somehow less valid than feeling good about yourself because you have just saved someone's life at risk of your own or solved an important problem.

Virtual reality gets us closer to Nozick's experience machine. Why do I feel better about making Germanic lyres in my basement than Whitesoul Helms in *World of Warcraft*? Why do I feel less comfortable about consuming many hours online fighting computer-generated monsters than about consuming a similar amount of time, also online, arguing with people about subjects of interest to me?

The problem is older than *World of Warcraft*. I know some very smart people who put substantial amounts of time and effort into playing games: chess, bridge, poker. As long as one views it as recreation, there is no problem. But what about someone who treats the game as his real life and whatever he does to earn money for food and rent as an annoying distraction? Something about that feels wrong to me, feels as though he would be leading a better life if he put the same talent and passion into building better houses or writing better computer programs. But I can't prove it; I might be wrong about values or he might be someone for whom the alternative I thought better was, for one reason or another, not an option.

Should Irrational Preferences Count?

The field of behavioral economics discussed in Chapter XXX[Beyond Rationality] deals with predictable patterns of behavior that appear inconsistent with rationality. My one contribution to the field is a chapter, "Economics and Evolutionary Psychology," in the book *Evolutionary Psychology and Economic Theory*; a [draft](#) is available on my web page. In it I try to show that some patterns of behavior which are puzzling in terms of the assumptions of economics make sense in terms of evolutionary psychology, can be explained as behavior that got hardwired into us because it increased an individual's reproductive success in the hunter-gatherer societies where our species spent most of its history.

Consider, as one example, the endowment effect, the observation that individuals value items that belong to them more than items that do not even if, as in the classic [Cornell coffee cup experiment](#), who owns what is the result of random chance. I explain this as a commitment strategy that acts to enforce property rights in a world without police and courts, the human elaboration of the territorial behavior observed in many animal species. If I am willing to fight harder to defend what is mine than to take what is yours, you are similarly willing, and both of us know it, then most of the time ownership will not have to be fought for.⁴

Economists define efficiency in terms of people getting what they want, not what they should want. Suppose you accept my explanation for the endowment effect, that it exists not because it serves the present interest of the individual but because it served the reproductive interest of other individuals long ago in a very different environment. Should you still take that preference as a given in evaluating economic institutions?

Why Risk Aversion Isn't

Many fields use technical terms that sound self-explanatory and aren't; many people believe they know what those terms mean and don't. Millions of people believe that they understand the Theory of Relativity, even if not the mathematical details. The theory says that everything is relative. Surely that is clear enough.

Clear but wrong; that is not what the theory says. One of the many things the Theory of Relativity tells us is that the speed of light is quite impossibly absolute, the same relative to you however fast you are moving.

Economics has similar problems with terms such as *efficiency* and *competition*.

And *risk aversion*. It sounds as though it means aversion to risk; one might expect a risk averse person to avoid dangerous hobbies, a risk preferring person to be drawn to them. It is not true. There is nothing in the definition of risk aversion that implies that a risk averse person is less likely to take up hang gliding or mountain climbing than a risk preferer.

The definition of risk aversion, as any good textbook that covers the subject will [explain](#), is that a risk averse person, faced with the choice between an uncertain set of monetary payments and a certain payment with the same expected value, will prefer the latter. That is a statement not about his taste for risk but about his taste for money.

To see why we would expect people to be risk averse, imagine that you are faced with two possible jobs. One pays you \$60,000/year. The other has equal odds of paying you \$20,000/year or \$100,000 year. The expected value of pay is the same for both, but we expect most people to prefer the former job, all else being equal. Why?

If you accept the risky alternative you are giving up dollars in the future where you lose the bet in order to get dollars in the future where you win the bet. You are giving up (probabilistic) dollars used to buy things you would get as your income increased from \$20,000 to \$60,000 in order to get (probabilistic) dollars to buy things you would get as it increased from \$60,000 to \$100,000. As your income increases you buy the more important things first, so we would expect the gain from getting a dollar at the high end to be less than the loss from losing one at the low.

⁴ I discuss this approach to enforcing rights in much greater detail in chapters [51](#) and [52](#) of the third edition of *The Machinery of Freedom*.

As this (entirely conventional) exposition shows, risk aversion is simply declining marginal utility of income. The fact that your marginal utility of income decreases as your income increases tells us nothing at all about how the marginal utility of other things changes as the amount you have of them changes, hence the fact that you are risk averse does not tell us what your attitude will be to risks that involve non-monetary payoffs. For example ...

Your doctor calls you into his office to give you some very bad news. You have been diagnosed with a disease that, if untreated, will kill you in fifteen years. There is an operation which, if it succeeds, will let you live thirty years — but half the time it kills the patient. You have a choice of a certainty of fifteen years or a fifty/fifty gamble between thirty and zero.

As it happens, the one thing in life you most want to do is to produce and bring up children. Thirty years is long enough to do that, fifteen is not. You grit your teeth and sign up for the operation. You may be, probably are, risk averse in dollars, which have decreasing marginal utility to you. But you are risk preferring in years of life, because years of life have increasing marginal utility to you; thirty years are worth more than twice as much as fifteen.

Risk preference is not about risk.

In Defense of Cardinal Utility

The economists of the late 19th treated utility as cardinal, quantitative, able to be represented by numbers. Early in the Twentieth Century John Hicks showed that everything then being done by treating utility as a number could be done by treating it as an ordering instead, a description of what you preferred to what but not by how much. About ten years later Von Neumann and Morgenstern demonstrated that there was something cardinal utility could do better than ordinal utility: incorporating preferences under uncertainty into the utility function. They showed that one could assign numerical utility values to outcomes in such a way that an individual choosing between lotteries, probability mixes of outcomes, would always prefer the lottery with the higher expected value of utility.

That is one reason to prefer cardinal utility, but not the only one. We observe other people's utility functions from the outside, our own from the inside. Seen from that perspective, it is clear to me that some of my preferences are much stronger than others — and in a way that fits at least roughly with Von Neumann's approach to utility theory. If I not only prefer an orange to an apple, I very much prefer it, I probably prefer even one chance in four of getting an orange (expected value = (utility of orange)/4) to a certainty of getting an apple (expected value = (utility of apple)). Cardinal utility, indeed Von Neumann utility, fits our subjective experience of our own preferences

A third reason to prefer cardinal utility is that it makes it easier to intuit economics. Marginal utility was a central element of the marginal revolution, the collection of intellectual breakthroughs that, at the end of the 19th century, created modern economics, a concept that proved useful to explain prices, economic equilibrium, and much else. If utility is only ordinal, an ordering not a quantity, there is no such thing as marginal utility.

Proving that something is true is not an adequate substitute for knowing why it is true; one of the problems with modern economics is the emphasis on formal mathematics to the neglect of economic intuition. One reason may have been the shift from thinking about economics the way Alfred Marshall did to thinking about it the way John Hicks did. The further your mental model is from what you understand, the more you have to rely on formal mathematics instead.

It is easier to understand declining marginal utility than the convexity to the origin of an indifference curve.

In Defense of Interpersonal Utility Comparisons

One of the common criticisms of utilitarianism is that we have no way of trading off utility gains to one person against utility losses to another, no way of knowing whether the net change is an increase or decrease in total utility. That is not a problem for much of economic theory, which can be done not only without interpersonal comparisons but without anything more than individual ordinal preferences. But it is a problem if you want some way of evaluating the overall effect of a change, such as imposing or removing a tariff, that makes some people better off and some worse.

In fact, we not only can make interpersonal utility comparisons we routinely do make them, even if not very well. A parent making decisions that affect his children is implicitly asking himself whether doing something one child wants to do and the other doesn't will increase the former's happiness more than it decreases the latter's. Someone deciding which friend to give a gift to is doing it in part on which he thinks will be made happier by it. I signal my feelings, including preferences, in facial expressions, voice tones, and the like. Others appear to do the same, giving me some idea of the strength as well as the ordering of their preferences and comparing to mine.

I cannot know another person's preferences with certainty but I have no serious doubt that the disutility to a random stranger of being tortured to death is greater than the disutility to me of stubbing my toe.

Interpersonal Utility in Economics

Alfred Marshall defined the value of an outcome to an individual, positive or negative, as the largest amount he would be willing to pay to get it or prevent it, defined an economic improvement as a change whose total value summed over everyone affected by the change, was positive. He offered the concept of an economic improvement as an imperfect proxy for a utility increase. Although a given amount of money might represent more utility for one person than another, which is why it is only an imperfect proxy, he argued that such differences would usually average out for changes that affected many people. Put in modern terminology, an economic improvement is an increase in economic efficiency and an outcome is efficient if it cannot be improved.

A Pareto improvement, a concept originated by Vilfredo Pareto, is a change that benefits at least one person and harms nobody, avoiding the need for any interpersonal comparison of amount of benefit and harm. An outcome is Pareto efficient if it cannot be improved. The problem with substituting the Pareto versions of "improvement" and "efficient" for the ones based on Marshall's approach is that almost no change affecting a significant number of people is a Pareto improvement, hence almost all outcomes are Pareto efficient.

Hicks and Kaldor tried to solve that problem with the concept of a potential Pareto improvement, a change that would be a Pareto improvement if combined with a suitable set of payments from people who gained to people who lost. If gainers gain more than losers lose, making the change a Marshall improvement, there should be some set of transfers that fully compensates the latter while leaving some gain for the former, making it a potential Pareto improvement, so the two concepts

are very nearly identical.⁵ Since the transfers are not actually made, a potential Pareto improvement is not an actual Pareto improvement — some people gain, some lose — so justifying it as a criterion for what changes are good or bad requires the same interpersonal utility comparison as Marshall's approach.

It just makes the fact less obvious.

In order for economists to conclude that abolishing a tariff or a minimum wage law or practically any other change is (or is not) good for the country, an improvement, they must be willing to bite the bullet, treat utility as interpersonally comparable.⁶

⁵ I describe a situation in which something is a Marshall improvement but not potential Pareto/Hicks-Kaldor improvement in "[Does Altruism Produce Efficient Outcomes? Marshall vs Kaldor.](#)" *Journal of Legal Studies*, 1987 Vol. XVII, (January 1988).

⁶ A much more detailed [explanation](#) of these ideas from my webbed *Price Theory*.