

Mill, *Utilitarianism*, chapter 2, pages 12-33. Blackboard Notes.

The rest of the chapter is organized as a series of objections to utilitarianism followed by Mill's replies. My interpretive and critical comments that go beyond straight exposition are enclosed in brackets, like this [].

Page 12: Objection: "Happiness cannot be the rational purpose of human life and action; because, in the first place, it is unattainable."

Mill's reply: Not so.

Page 15: Objection: Happiness is not the rational end and purpose of human life and action; virtue is a better end or goal than happiness.

Mill's reply: Virtue is morally desirable character. Individual virtues are morally desirable character traits like courage, fortitude, temperance, honesty, etc. What character traits are morally desirable? The ones that do most to promote total human happiness. An overall virtuous character, rightly conceived, is a set of traits that produces more total human happiness than alternative feasible sets of traits. Virtue is desirable as a means to happiness, not desirable intrinsically, for its own sake.

Page 16: If a person makes her happiness the goal of her life, that decreases her chances of becoming happy. Pursuing happiness is self-defeating (the paradox of happiness).

Mill's reply: It's possible that directly aiming at one's happiness is self-defeating as the objection alleges. Just suppose this is true. Then to organize one's life to maximize one's chances of happiness one must follow an indirect strategy—not aiming directly at happiness but learning to care about various things for their own sake, caring about which will tend to make one happy. Maybe the same goes for pursuing the happiness of other people—to be successful in this pursuit one must follow an indirect rather than a direct strategy. If so, pursuing happiness (for self or others) is not necessarily self-defeating.

Page 16: Utilitarianism tells each person she ought above all to pursue her own happiness. Doing that would be selfish.

Mill's reply: "the happiness which forms the utilitarian standard of what is right in conduct is not the agent's own happiness but that of all concerned."

Page 17: Objection: Utilitarianism sets too high a standard of conduct for ordinary people. "It is expecting too much to require that people shall always act from the inducement of promoting the general interest of society."

Mill's reply: "But this is to mistake the very meaning of a standard of morals and to confound the rule of action with the motive of it." Utilitarianism is a criterion of morally right and wrong action, not necessarily a practical decision-making guide. Being motivated to conform always to utilitarian principle is not necessarily the motivation it is best to have by the utility-maximizing standard. Mill holds that our actions mainly affect ourselves and those near and dear to us, and that for the most part, the agent's concern just for this narrow circle will suffice to induce right conduct. [Example: Suppose that in particular circumstances, having a party is the right act, the act singled out as right by utilitarian principle. But it might well be that if one were motivated by the desire to do whatever is right, or whatever maximizes total utility, those motives would operate as a wet blanket, dampening otherwise available happiness. Perhaps being spontaneously motivated by the desire to have some fun will lead to having a party that is more enjoyable for all involved than being motivated by moral principle. Another example: Suppose Tom buys his mother a gift for her birthday, and when she says that was a nice thing to do, he says, "Yes, I calculated that buying you a present was, among the available alternatives, the act that would most maximize human happiness on the whole." Or perhaps even without saying anything his surpassingly high-minded motivation expresses itself involuntarily and is detected by his mother, who is left cold. She had hoped Tom was buying her a present because he loved her.]

[What motives should we cultivate in people, according to utilitarianism? One should according to utilitarianism always do whatever will be happiness-maximizing. When one's actions affect character and motive formation, one should act in ways that will be happiness-maximizing. Suppose one has a child and could raise her to be either like Mary or Martha, (cf. the New Testament characters). Martha always does the morally right thing according to utilitarian principle, but her motivation dampens happiness. Mary is carefree and charming. Her strongest motivations are to be sympathetic and kind and fun-loving. These motivations lead her to do some morally wrong acts, but overall her life generates more happiness for people than Martha's.

So here's the choice:

Choice	Results of Choice	
Form one's child to be EITHER	Does the morally right act (happiness-maximizing act)?	Causes overall happiness for people?
Martha	More	Less
OR		
Mary	Less	More

Act utilitarianism says one morally ought to form one's child to be a Mary type and not a Martha type if the results are as just stated. Mary does some acts that are wrong by the utilitarian standard but is a fountain of happiness for self and others. Martha always does the right thing by utilitarian standards but with this motivation, causes people to uncomfortable and unhappy.]

Page 21: Objection: Utilitarianism is a godless doctrine.

Mill's reply: Whether or not there is a conflict between religion and utilitarian ethics depends on one's religious views. If one conceives of a God who loves humanity, belief in such a God may well be fully compatible with affirming an ethics that bids us to promote the greatest happiness of the greatest number. Also, belief in God is compatible with the belief that God intends that humans figure out how to live mainly by reflectively deciding on ethical principles rather than by following mechanically the instructions laid down in some sacred book thought to contain divine revelation.

Page 22: Objection: Utilitarianism instructs people to do what is expedient rather than to stand fast by principle. In the name of expediency one might lie, steal, cheat to maximize overall human happiness.

Mill's reply: In ordinary speech, following the path of expediency rather than principle usually involves pursuing short-term gain at the expense of greater long-term harm or one's personal gain at the cost of greater losses suffered by others. In this sense, acting in ways that are expedient is opposed to acting in ways that are right according to utilitarian principle. Regarding the thought that utilitarianism leaves it an open question whether an individual ought to obey or violate common-sense moral rules such as "Tell the Truth" and "Keep your promises," Mill replies that following these rules strictly in most circumstances is following the course of conduct most likely to be utility-maximizing. But no common-sense moral rule should be obeyed come what may, regardless of consequences. In unusual circumstances acting against even a deeply entrenched moral rule can be utility-maximizing, and in those cases should be done.

Page 23: Objection: "there is not time, previous to action, for calculating and weighing the effects of any line of conduct on the general happiness."

Mill's reply: In such circumstances, one should follow common-sense moral rules, which summarize lots of human experience, and tend to guide us toward actions that promote general happiness and away from actions that tend to dampen it. Also, one can cultivate habits and train individual character, so that people become disposed to act in ways that are happiness-promoting. "Whatever we adopt as the fundamental principle of morality, we require subordinate principles to apply it by." Adopting utilitarianism as the fundamental principle, we need also subordinate principles—moral rules that we train ourselves and other to follow because following them will for various reasons produce better results as assessed by the fundamental principle than we would get if we directly applied the fundamental principle case by case to decide what is due.

UTILITARIANISM AND CONFORMITY TO COMMON-SENSE RULES.

Objection: The moral person, the person with integrity, stands fast by important moral rules including tell the truth, keep your promises, be loyal to your friends, refrain from attacking innocent nonthreatening persons, respect the private property of others, be faithful in love, and so on. In contrast, utilitarianism gives no moral standing to such rules. Utilitarianism says one morally ought always to do whatever will maximize happiness (the general welfare), or in other words lie or refrain from lying, kill the innocent or not, depending on what would in one's particular circumstances produce the best outcome (most total happiness).

Most of what Mill says in chapter 2 on this issue is consistent with **act utilitarianism (roughly, the righter/wronger test)**. **One ought always to do whatever is utility-maximizing**, and in the actual and likely circumstances of human life, generally following the moral rules is what would maximize utility (happiness) compared to alternatives. But when that is not so, one should violate not follow the rules. To give the rules more weight than that would be to fall into unjustifiable rule worship.

Sometimes Mill suggests another proposal, different from and opposed to act utilitarianism. This would be some version of rule utilitarianism. According to rule utilitarianism,, a two-step process fixes what is morally right to do. Step one: Determine the set of rules that is utility-maximizing. Step two: Obey that code of rules (justifiable in terms of the utility of the rules).

There are various versions of rule utilitarianism. Mill suggests one version in a passage on page 19: “In the case of abstinences indeed—of things which people forbear to do from moral considerations, though the consequences in the particular case might be beneficial—it would be unworthy of an intelligent agent not to be consciously aware that the action is of a class which, if practiced generally, would be generally injurious, and that this is the ground of the obligation to abstain from it.” Here Mill says what makes an act right or wrong is not the quality of its actual consequences (compared to those of feasible alternatives) but rather the hypothetical consequences that would be generated if everyone made it a rule to behave the same way in relevantly similar circumstances. This suggests **ideal rule utilitarianism: One morally ought always to act in conformity with that ideal code of rules, conformity to which by everyone would maximize utility.**

This version of rule utilitarianism attracts the objection that the fact that a code of rules would produce ideal consequences if everyone were to obey it does not seem to settle the question what one should do in the many actual situations in which others are not conforming to this hypothetically ideal code. Pacifism (always refrain from using violence and threat of violence to settle conflicts among persons) might well produce great consequences if everyone were to conform to it but bad consequences in the many actual and likely circumstances in which other people are not conforming to pacifism whatever one does.

Apart from the lone page 19 passage, the rest of chapter 2 seems consistent with an act utilitarian approach to social rules. [The key is to note that if everyone were maximally well informed about the circumstances in which she acts, ideally smart at figuring what act would maximize utility in her actual circumstances, and ideally disposed to do whatever act utilitarianism requires, then each person should always use act utilitarianism as a decision procedure for choosing what to do. But we actual humans are often not well informed, not ideally smart at figuring out what act of those we might choose would produce most utility in the circumstances as we (perhaps incorrectly) understand them, and disposed to be selfish and favor those near and dear to us, not disposed to act according to impartial utilitarian principle. For people like us, actual humans, we would usually do better at fulfilling the act utilitarian principle if we did not directly apply this principle to our choice of what to do but instead followed simple social rules like “Tell the Truth” and “Do no harm” and “Keep your promises.” Call the ideal agent who is perfectly well informed, maximally intelligent, and ideally disposed to follow impartial act utilitarian principles an “archangel.” When one faces a particular decision problem, the more it is the case that one resembles the archangel, the more one should set aside moral rules (or what Mill calls “subordinate principles”) and directly apply the act utilitarian test to decide what to do, and the less one resembles the archangel, the more one should be disposed to forego direct act utilitarian calculation and just follow whatever the going moral rules say one should do in one’s situation. Either way, the criterion or test of the moral rightness of what one does, on this view, is act utilitarianism. At the level of fundamental moral principles, the theoretical determiners of moral right and wrong, act utilitarianism (on this view) is correct, but on the level of practical decision making, the best guides to follow (best by the act utilitarian standard) may be the going moral rules. Following this same train of thought, one might hold it is right, according to act utilitarianism, to train people to accept and follow simple moral rules rather than to apply act utilitarian calculation case by case to the choice of what to do. (The ideas in this paragraph are borrowed from R. M. Hare, not a course author.) When one follows the going social/moral rules, the consequences will be better or worse, according to utilitarian assessment, depending on the quality of the rules actually entrenched in one’s society. Recall Huck Finn.]

Rawls: A third alternative. John Rawls argues that utilitarianism can be stated in a way that makes it not vulnerable to a standard objection if one keeps in mind the distinctions between the summary and practice conceptions of rules and between justifying an act that falls under a practice and one that does not. He uses the social practices of punishment and promising to illustrate his points. The standard objection is that utilitarianism holds that promise breaking is justified in any case in which breaking a promise produces even a slight net gain of utility on the whole. Also, utilitarianism justifies punishing a person of a crime he did not commit if doing so in the particular circumstances would produce even a slight net gain of utility. In each case, the objection runs, utilitarianism fails adequately to register the obligation to keep a promise one has made and the duty not to punish the innocent as considerations that should affect what one ought to do.

Rawls contends that the discussion fails adequately to register the fact that punishing and promising are social practices. A social practice is an activity constituted by rules in the sense that the rules define and make possible behavior that could not exist absent the rules. For example, the rules of baseball define what it is to play baseball and specify, for example, that after three strikes, you’re out. Absent the rules defining the practice there could be no such thing as a pitcher striking out a batter.

We should think of the justification of punishing and promising as proceeding in two steps for the utilitarian. If one asks, in a setting where the rules specify what one should do, how should one act, the answer is given by the rules. In a criminal trial, the judge should follow the rules of criminal procedure. So if one asks, why put Smith, the accused, in prison, the answer is that he was convicted of an offense after a fair trial and has been sentenced to four years of imprisonment. In a similar way, if one asks, why should I return this necklace to Mary now, if one has promised to do

so, the answer refers to the informal rules of the promising institution. One should return the necklace because that is what one promised to do and promising involves undertaking an obligation to do what one has promised. But we can ask another sort of question. We can ask, why have a social practice with these rules. According to the utilitarian, the answer is that if the practice is justified, then its rules are such that the operation of the practice works to advance utility and alteration or abolition of the practice would not result in greater utility but less.

According to Rawls, there is a confusion if one supposes a utilitarian calculation might yield the verdict that one should pay no heed to the promise one has made or that the judge in a criminal trial should rig the proceeding to procure the judicial punishment of the innocent. By what authority would the judge act in this way? The practice does not give such discretion to those who act under it and moreover it would not be a good strategy to promote utility to alter these practices to give agents a wide discretion to suspend the rules and do whatever they think is utility-maximizing in the circumstances. One has to envisage an institution with rules that allow one who has made a promise to give it no weight in further decisions about what to do or that allow a criminal prosecutor to falsify evidence to secure the result she deems best on the whole. Building such exceptions to the normal rules into the practice and giving such wide discretion to those regulated by the practice would be an outstandingly bad utilitarian strategy. But if one cannot devise a practice that allows “punishing” the innocent and that would be justifiable on utilitarian grounds, then in the course of operation of the practice of criminal punishment it cannot be a live issue whether one should bring about “punishment” of the innocent. What one should do when acting within a practice is set by the rules of the practice. In baseball, it would be a bad joke if a batter on a particular occasion proposed to the umpire that she should have four strikes because making this allowance would produce more happiness on the whole.

Of course, our ordinary understanding of a promise does not hold that a promise may never be broken no matter what. The ordinary practice of promising includes a vague escape clause, so that when one promises, a proviso is implicit, to the effect that if the consequences of keeping the promise turned out to be unexpectedly extremely bad, the promise would no longer be binding. But this type of exception is built into the rules that define the practice. In a somewhat similar way, the rules of criminal procedure do not guarantee that no person accused of a crime and put on trial will never be found guilty and punished by a fair proceeding if she is innocent. The rules of a fair trial cannot guarantee perfect outcomes. But the law of criminal procedure does unequivocally forbid a judge, jury, or prosecutor deliberately and knowingly to seek to bring about the judicial punishment of someone for an offense who is innocent of that offense.

Rawls distinguishes what he calls the summary and practice conception of rules. Both conceptions make sense. According to the summary conception, a rule is a rule of thumb—a guide to decision making that in effect summarizes the results of past experience and gives guidance about how to proceed to get desired outcomes. The rules of gardening such as “plant in the spring,” “weed your garden vigorously,” and “harvest just at the peak of ripeness of the crop” are examples of summary rules. They guide behavior that is itself not defined or constituted by the rules. In contrast, when a rule is a practice rule, it established or defines kinds of behavior that could not exist absent the rule. In this sense the rules are logically prior to the behavior the rules regulate. One could not strike out a batter—there would be no such thing—except that the rules of baseball specify what counts as a batter, a pitcher, a strike, and so on.

According to Rawls, both many utilitarians and their critics make the mistake of failing to notice the existence and significance of rules of practices. They discuss utilitarianism and its implications for right conduct and right policy as though all rules were summary rules. But this is a fatal error. If one is dealing with summary rules, the rules have no force as reasons except as guidance about what behavior is likely to get good results. It would be objectionable rule worship to appeal to the summary rules as authoritative determiners of what one ought to do in circumstances where it is acknowledged following these rules will do no good. But it is not mere rule worship but acknowledgement of the logical distinction between summary and practice rules to deny that utilitarianism properly understood can recommend violating the rules of a practice, for example, by procuring deliberately the judicial punishment of the innocent. (To avoid the irrelevant concern that by definition punishment is imposition of hard treatment on one who is guilty of an offense, Rawls considers whether utilitarianism might go wrong by favoring in some circumstances telishment of the innocent, where telishment is just like punishment but may sometimes be imposed on the innocent. Rawls says that one would have to formulate a set of rules that define telishment to assess it from a utilitarian standpoint, and once one makes the assessment at the level of practice rules, it is obvious that no such practice as telishment could be justified as utility-maximizing.)

By the way, Rawls in this essay is not asserting that utilitarianism is justifiable all things considered. In the essay he does not commit himself one way or another as to the overall acceptability of the utilitarian doctrine. His point is that one standard objection against utilitarianism can be seen to misfire once one takes account of the distinction between summary and practice rules and properly considers what the utilitarian should say about the justification of a practice and the justification of behavior that is internal to the practice.

Objections to Rawls’s analysis.

1. As stated, Rawls's view is incomplete. He says that according to utilitarianism the rules of practices should be set so as to maximize utility and individuals engaged in practices should follow their rules. As it stands, this says that what an individual should do when a practice rule applies depends on the actual rule in force, and assumes it has been set so it has a utilitarian justification. This view is incomplete in that it does not tell individuals what to do when they are engaged in practices whose rules that are not optimal from a utilitarian standpoint.

2. More important, Rawls seems to overstate the reason-giving force of an applicable practice rule in determining what a person ought to do. Suppose I am a judge in a jurisdiction shaken by a terrible crime wave. Exemplary punishment imposed swiftly on some prominent accused persons who are universally believed to be guilty as charged will predictably lower the crime rate and save lives of innocent crime victims. It turns out that these prominent accused persons are not in fact guilty, but the evidence that shows this is known only to the prosecuting attorney and the judge and they can conspire to keep it secret in violation of the criminal code. The judge then can choose to follow the practice rules of criminal procedure or violate them and produce a net saving of innocent lives. The act utilitarian (I submit) will then say that if the overall consequences of violating the rules are utilitarianly better than the consequences overall of abiding by the rules, the judge ought to violate the rules. In this case, it is true that the judge's behavior could not be what it is absent the rules. But this will be true whether she conforms to the rules or violates them. The fact that one's behavior is only describable as "violating the rules of criminal procedure" or as "conforming to the rules of criminal procedure" because the practice rules exist does not tell the act utilitarian judge whether she ought to abide or conform. From an act utilitarian standpoint, Rawls goes astray in supposing that the mere existence of a practice rule has magical force in determining what one ought to do. This is from the act utilitarian standpoint a species of rule worship. Understanding the logical distinction between summary and practice rules does not have the implications for what one should do that Rawls supposes.

One should understand the difference between establishing practice rules and doing what the rules prescribe in a particular case. But the act utilitarian judge who secretly conspires to bring about the punishment of the innocent to bring about the greater good need not be conceptually confused in any way. In fact it would be a confusion to suppose that anyone who favors breaking the rules of a practice on a particular occasion must somehow be recommending an alteration in the rules of the practice. It may be that the legislators establishing a practice find the rules that are best from a utilitarian standpoint and enact just those perfect utilitarian rules. It still might be that on some particular occasion, one can gain extra utility by violating the rules of a practice even if the rules are impeccable from a utilitarian standpoint. To fail to appreciate this possibility is to fail to understand the idea of a practice rule.

Rawls's baseball example that seems to support his case actually undermines his position if you think about it. Suppose we are playing sandlot baseball and Sally proposes that it would be best (happiness-maximizing) on the whole to give her little brother Sammy four strikes. Sometimes in this situation the other players agree and Sammy is given four strikes (or perhaps one just surreptitiously lets Sammy have an extra chance at bat without anyone else noticing this anomaly). The fact that this conflicts with the rules of baseball does not show it is unjustified or could not be justified. No doubt if players were always proposing to renegotiate or suspend the ordinary rules the game could not proceed. But this precludes neither occasional relaxation or suspension of the rules nor cheating against them. There is something to be said against messing with the rules of practices but this does not begin to show it is never justified according to act utilitarianism. What is true of informal ("sandlot") practices is true of all others including formal practices in which playing the game by the rules is viewed rightly as important by spectators and participants. Still, there may be exceptions that an act utilitarian should recognize, and the mere fact that some behavior is internal to practices does not gainsay this possibility.

Question: what's the upshot of this discussion? You might consider whether Rawls's argument can be defended against the criticisms just noted. Even if it can't be defended, this would not show act utilitarianism is correct, but only that Rawls's criticism of act utilitarians and their critics (that both ignore the importance of the distinction between the practice and summary conceptions of rules) misfires. The objection against utilitarianism that it is unfair to punish the innocent even if it would maximize human happiness on the whole to do so might be a good objection for all that Rawls has said. Mill takes up this question in chapter 5.