

Rule utilitarianism¹

Act utilitarianism judges the rightness or wrong of an act by its consequences. It claims

1. Actions are morally right or wrong depending on their consequences and nothing else. An act is right if it maximises what is good.
2. The only thing that is good is happiness.
3. No one's happiness counts more than anyone else's.

If we put (1) and (2) together, we see that the theory claims that an action is right if it *maximises* happiness, i.e. if it leads to the greatest happiness of all those it affects. Otherwise, the action is wrong.

Rule utilitarianism agrees with claims (2) and (3), but disagrees with (1). It claims that an action is right if, and only if, it complies with those rules which, if everybody followed them, would lead to the greatest happiness (compared to any other set of rules). In other words, rule utilitarians do not look at the consequences of individual actions, but at the consequences of people following rules. A rule is morally right if following it leads to greater happiness than following an alternative rule. An action is morally right if it complies with a rule that is morally right.

SMART ON RULE UTILITARIANISM

In *Utilitarianism*, J.J.C. Smart objects that rule utilitarianism amounts to 'rule worship'. The point of the rules is to bring about the greatest happiness. If there is a situation in which breaking the rule will lead to more happiness than following the rule, what reason could we have for following the rule? If I know e.g. that lying in a particular situation will produce more happiness than telling the truth, it seems pointless to tell the truth, causing unhappiness. The whole point of the rule was to bring about happiness, so there should be an exception to the rule in this case. The alternative to 'everyone obeys the rule' is not 'no one obeys the rule'. Clearly, the best course of action is for someone people to break the rule sometimes, in which situations in which breaking the rule leads to more happiness than following the rule.

Rule utilitarians could respond by saying that we should amend the rule to allow the exception in such cases. For example, 'Don't lie' should become 'Don't lie unless telling the truth will hurt someone's feelings'. However, life is complicated. Whenever a particular action causes more happiness by breaking the rule than by following it, we should do that action. Smart points out that, if we keep amending

¹ This handout is based on material from Lacewing, M. (2017) *Philosophy for AS and A Level: Epistemology and Moral Philosophy* (London: Routledge), Ch. 3, pp. 240-5

the rules like this, there is no difference between what an act utilitarian would recommend and what a rule utilitarian would recommend. Furthermore, and more seriously for rule utilitarianism, if we try to add all the possible amendments to the rules in order to make acting on them always produce the greatest happiness, we will end up with just one rule, namely to maximise the greatest happiness. Because there are so many situations in which breaking a rule might lead to greater happiness than following it, no other rule can be certain of maximising happiness. And so rule utilitarianism collapses into act utilitarianism. It isn't a distinct theory at all.

RULE UTILITARIANISM DEVELOPED

To answer this objection, rule utilitarians can do two things. First, they could provide some reason for following rules even when doing so doesn't maximise happiness. Second, they could argue that despite recommending that we follow rules when doing so doesn't maximise happiness, it is still a better theory than act utilitarianism. In fact, rule utilitarians have provided arguments that support both points at once.

Their first argument is that morality should be understood as a set of rules. It needs to provide general guidance over the long term, a way of thinking that people can remember and rely on. So rather than considering actions individually, we need to take the bigger picture in thinking about right and wrong. This is just a reflection on what morality must be, psychologically speaking. Since we are to have rules, then the aim of these rules should be to maximise happiness. And so actions are right when they follow a rule that maximises happiness overall - even when the action itself doesn't maximise happiness in this particular situation.

Second, rule utilitarians argue that their theory has a number of advantages over act utilitarianism.

1. Problem with calculation: Act utilitarianism requires that we work out the consequences of an action for human happiness. But is this realistic? How can we know or work out the consequences of an action, to discover whether it maximises happiness or not? Surely this will be too difficult and too time-consuming for us to do.

But according to rule utilitarianism, we don't have to work out the consequences of each act in turn to see if it is right. We need to work out which rules create the greatest happiness, but we only need to do this once, and we can do it together. (This is what Mill says human beings have done over time, giving us our customary moral rules.) Rule utilitarianism gives rules a formal place in its theory of whether an action is right.

2. Fairness, liberty and rights: Act utilitarianism does not rule out any type of action as immoral. There are no constraints on pursuing the greatest happiness. For example, if torturing a child produces the greatest happiness, then it is right to torture a child. Suppose a group of child abusers only find and torture abandoned children. Only the child suffers pain (no one else knows about their activities). But they derive a great deal of happiness. So more happiness is

produced by torturing the child than not, so it is morally right. This is clearly the wrong answer. More generally, act utilitarianism, we may object, doesn't respect individual rights or liberty, because it doesn't recognise any restrictions on actions that create the greatest happiness.

By contrast, rule utilitarianism can plausibly argue that a rule forbidding torture of children will clearly cause more happiness if everyone followed it than a rule allowing torture of children. So it is wrong to torture children. More generally, individuals have rights, which are rules, because if people have to follow these rules (respect people's rights), that leads to the greatest happiness. Rules requiring fairness and justice will produce greater happiness in the long run than rules that do not.

3. **Partiality:** Many of the things that we do to make people happy are aimed at specific other people, viz. our family and friends. But act utilitarianism argues that in our decisions, we need to consider the greatest happiness that our actions could create, and this requires us to consider the happiness of each person equally. In other words, we should be impartial. If act utilitarianism is right, it seems we should spend much less time with the particular people we love and more time helping people who need help.

However, rule utilitarians can argue that a rule that allows partiality to ourselves, our family and friends will create more happiness than a rule that requires us to be impartial all the time. This explains the moral importance of partial relationships - they are necessary to happiness. Of course, we shouldn't be completely partial. We still need to consider the general happiness, but we only need to act in such a way that, if everyone acted like that, would promote the greatest happiness. For example, in the case of charity, I only need to give as much to charity as would be a 'fair share' of the amount needed to really help other people. This combination of partiality and impartiality respects both our natural inclinations and the demands of morality.

4. **Moral integrity:** Having integrity involves acting according to your own values, sticking to them especially in the face of temptation or other situations that would make it easier to do something you consider wrong. Just as act utilitarianism appears to require us to set aside our partiality, it can also seem to require that we set our moral values in order to maximise happiness. In other words, utilitarianism attacks our moral integrity.

The rule utilitarian may provide a similar argument for integrity as partiality. The best rule, perhaps, will be one that allows exceptions to other rules, i.e. a rule that allows people not to follow other rules if they conflict with one's integrity.

5. **Intentions:** Act utilitarianism claims that an action is right if it leads to the greatest happiness. It does not, therefore, recognise the moral value of our intentions in acting as we do. But whether someone intends to harm us or not - whether or not they do harm us - makes a big difference to how we respond to their action.

Rule utilitarianism can explain the importance of intentions by thinking about what it is to for an action to comply with a rule. It is not just a matter of what its consequences are. For example, a person kills someone else. But was this murder or manslaughter or something else again? If the person intended to kill someone, and that is what they wanted to bring about, that is very different than if the killing was accidental or if the person was only intending to defend themselves against an attack. A rule prohibiting murder is not a rule prohibiting self-defence, even if self-defence may lead to death. So whether an action complies with a rule or not depends, at least in part, on the individual's intention in performing the act. Two actions may have the same consequences, and yet one is right and one is wrong, because they are different types of action and fall under different moral rules, because the person's intention in each was different. And so rule utilitarianism can recognise that the thought does count.

These developments in the theory enable the rule utilitarian to respond to Smart's objection. If people try to follow act utilitarianism, this will lead to less happiness in the long term. For instance, people will make mistakes in working out what the consequences of their action for happiness will be. They will no longer feel secure in their rights and there will be pressure on personal relationships to be given up in favour of impartiality. And we will fail to make important distinctions between different types of action. So we may have to give up some happiness here and now to be more certain of the greatest happiness overall. Following rules provides this certainty.

All this provides reasons to follow rules even when doing so doesn't lead to the greatest happiness on occasion. An act is not made right by maximising happiness but by being in accordance in a rule. If there is a conflict between following the rule and maximising happiness (e.g. telling a lie when 'no harm is done'), there should be no temptation for the utilitarian to say that we should break the rule. Even if the act does maximise happiness, this doesn't make it right.

OBJECTIONS

We may question whether the replies that rule utilitarianism gives above are persuasive.

First, we can object that rule utilitarianism still fails to understand what is important about partiality. Rule utilitarianism permits partiality and it encourages it insofar as partiality contributes to happiness. But partiality is good, according to rule utilitarianism, because it contributes to the general happiness. The importance of the individual, and making someone happy just for their sake, still seems to be missing. For instance, if I form friendships because this maximises happiness, does that respect and value my friends for themselves, as the particular people they are? Consider this example from Bernard Williams. Suppose a man is in a boating accident with both his wife and a stranger. Neither can swim, and he can only rescue one. We might think that he should simply rescue his wife. But if he thinks, 'Rescuing my wife is in accordance with a rule that will lead to greater happiness than a rule that requires me to rescue the stranger', this seems to miss the particular importance that being married has, including its moral importance.

Second, a rule that protects integrity could prove very problematic. If what someone values goes strongly against promoting the greatest happiness, then morality itself may seem in conflict with their integrity. Is it possible, for instance, that someone finds meaning in their life through making as much money as possible, without constraints? The rule utilitarian must insist that people's values, and so their integrity, are broadly in line with what the set of rules that maximises happiness generally requires. Put another way, morality itself probably needs to be part of people's commitments for rule utilitarianism to respect their integrity. But now, is this an objection any more? Or simply a recognition of the legitimate demands that morality makes upon us?

Rule utilitarianism also faces an objection based on its theory of the good. It agrees with act utilitarianism that happiness is the only good. But is this right? Suppose morality is a set of rules. Should that set of rules aim to maximise happiness, or are there other important values, such as telling the truth, that matter independent of the pleasure they bring (or preferences they satisfy)?

A new objection that rule utilitarianism faces, which act utilitarianism doesn't, is whether all of morality can be summed up by rules. Isn't life too complicated for this? If so, we will need a different theory to explain what the right thing to do is when there are no rules that apply. The rule utilitarian has a simple answer to this objection, though we can question whether it is adequate. One of the rules is 'When no other rules apply, do that action that maximises happiness'.