

140: Matters of life and death – 4 – Utilitarianism and Justice

1 The problem of justice

Many of the problems with Utilitarianism relate to its normative consequences.

Allegedly, it tells us to do what we think we ought not to do, or not to do what we think we should. (Remember the method of reflective equilibrium.)

(a) – **punishment.** We may be obliged to seek the punishment of the innocent whenever doing so would create greater utility. The example of the Sheriff and the lynching party.

But this may undermine the rule of law, weaken the Sheriff's own respect for the law, and so reduce the utilitarian benefits of punishment in the future.

The Atlanta murders example. Since no one will ever know, there will be no weakening of respect for the law. How to explain the intuition that it would be wrong to frame the innocent man?

(b) – **slavery.** What matters is the *total quantity* of utility produced, irrespective of its distribution. So we might be forced to condone a system of slavery if the numbers of slaves were few, and any alternative system would produce less utility overall?

One response: in the actual world, real instances of slavery always produce more suffering and less good than would some viable alternative. Examples where this isn't so are merely imaginary.

How can this reply be relevant? A false hypothetical ("If circumstances were such-and-such, then slavery would be right") entailed by a theory can still falsify that theory!

2 Rule utilitarianism

Some Utilitarians have attempted to avoid the above difficulties (and so achieve reflective equilibrium) by moving to a new, restricted, version of the theory. Distinguish between–

act-utilitarianism – each individual action is judged by its consequences (either *actual* utility, or *expected*, or *reasonably expectable*). This is the form of the theory discussed so far.

rule-utilitarianism – individual actions are judged by rules, and the rules are judged by their utility. Direct calculations of utility are only to be made where there *is no* rule covering the case, or where rules *conflict*.

So we ought to reject a system of limited but hugely utility-producing slavery because it conflicts with the rule against slavery, and we should refuse to frame the innocent man because it conflicts with the rule against punishing the innocent, and we should refuse to break the desert island promise because it conflicts with the rule enjoining promise-keeping, etc.

But if we ask why these rules are *good* rules, the answer is that they are generally productive of utility. So we can keep all of the advantages of utilitarianism without any of the pain?

3 Reducing rule utilitarianism to act utilitarianism

An argument to show that the two must come down to the same thing, and can never actually disagree.

So if justice is a problem for act-utilitarianism, it must still be a problem for rule-utilitarianism.

Suppose **rule R forbidding acts of type A** – e.g. “don't punish the innocent”. The justification of R will be: **acts a_1 - a_{200} (all of type A) gave less utility than alternatives.**

Now consider case a_{201} where *act-utilitarianism* says one *should* break the rule – e.g. where an innocent is generally believed guilty, and punishing would end rioting.

Can a rule-utilitarian disagree about this case? There must be something about the circumstances that makes a_{201} different – e.g. that the breach of R will not be known by anyone but the agent. Call these "circumstances C".

Then we have **rule R* = don't do A except in C**. This will be *better justified* than R, since it covers all of the cases a_1 - a_{200} , and also case a_{201} . It would be unmotivated *rule-worship* for the rule-

utilitarian to insist on keeping to R.

The argument generalizes. Wherever there is an apparent conflict between act and rule-utilitarianism, this only shows that the latter has got hold of the wrong rule!

4 **Two levels of moral thinking** (R.M. Hare 'What is wrong with slavery?' in Singer ed.)

Hare's reply to the problem of slavery: distinguish the level of **theoretical reflection** from which, as utilitarians, we have to concede that if the right circumstances *were* to occur then slavery would be right; from the level of **day-to-day moral reasoning**, which engages settled attitudes, principles, and dispositions of thought and feeling. For we cannot be calculating consequences all the time, and human action is often the product of such settled dispositions and qualities of character.

When we think theoretically, we are (normally) thinking about what principles and attitudes we *should* develop in ourselves and others *in the world as it actually is*. And in the actual world, a deep abhorrence of slavery is the right attitude to have.

So: when we consider a hypothetical example of slavery, these attitudes are, very naturally, engaged – we feel abhorrence. Hence the intuition that slavery would still be wrong here, despite the overall greater utility. But this intuition is strictly false, for a utilitarian. The intuition is then explained by *explaining it away*.

But: how can an Utilitarian capture the intuition that it is wrong of the prosecutor to frame the innocent man? For this example need *not* be highly imaginary. Many real dilemmas of this sort may actually occur.

Is the distinction between judging *acts* and judging *agents* useful here? Maybe the judgment that framing is wrong is really a judgment of the agent? Maybe the judgment is that they are shown not to have the attitude towards the law which (from the theoretical level) we can see that they *ought* to have? **Note:** we surely want law-officers who won't even *consider* breaking the law, since they may otherwise be tempted into doing so when they should not.

But this still leaves in place the claim that framing was the right *thing to do* (a judgment of the act), even though the *agent* was wrong to do it. Is there a version of utilitarianism which can avoid this consequence?

5 **Quality-of-character Utilitarianism** (See J S Mill *Utilitarianism*).

Proposal – acts are to be judged by the qualities of character displayed, and only the latter are to be judged by their utility. (Formally similar to rule-utilitarianism, with a two-tier justification.) Recall the point of judging acts is to give guide for the future. But we won't become or remain people with the qualities of character we want unless we perform acts expressive of them.

The idea of *virtue* – dispositions of thought, feeling, or action (e.g. courage, loyalty, generosity) that often essentially involve *no calculation*. On utilitarian grounds we should try to become the sort of people who often (in our daily practical lives) do not engage in calculation at all. (E.g. fidelity, loyalty.)

So we can now say, not only that *the prosecutor* was wrong to frame the innocent man, but also that he *acted wrongly*. (Similarly for promise-breaking, slavery, etc.) So all our intuitions are captured?

Is there a risk of a reduction to act-utilitarianism again? But a disposition towards R*, rather than R, may be *psychologically unstable*, in real danger of degenerating into a bad state of character. Qualities of character may be *necessarily blunt instruments* from the point of view of maximizing utility.

⇒ It is, at least, *not easy* to get a definite refutation of utilitarianism on grounds of injustice. (But consider the example of the drug in limited supply, small quantities of which can cure acne, but the entire quantity of which will save a life. Carruthers, *Animals Issue* ch.2.)