

Options in Practical Rationality

According to many theories of practical rationality, including classical decision theory, you ought to do whatever option is ranked best, where an option's ranking is a function of its possible consequences. I want to solve a puzzle for such theories. The puzzle is that there seem to be cases where there is no suitable set of options. My solution is that your options consist of making true certain counterfactuals, namely, those of the form *if you could do ϕ , then you would do ϕ* .

My focus here is on the *subjective ought* – that is, the ought that is sensitive to your beliefs and desires – where this includes the subjective *moral ought* and the subjective *prudential ought*. The puzzle starts with the observation that consequentialist rankings of options – that is, rankings of options according to their believed consequences – ignore the agent's doubts about being able to do it.

The problem is that on a naïve conception of options this lands consequentialism in trouble. For instance, suppose that you are considering whether to save ten miners in shaft A or nine miners in shaft B (you can't do both). And suppose that you think it's very unlikely that you're able to save the ten because shaft A is far away and you won't be able to get there in time; in contrast, shaft B is very nearby and it'll be easy to save the nine. Naively, your options are *save the ten* and *save the nine*. In ranking these options, consequentialism looks only at your beliefs about the consequences of saving the ten and saving the nine. And you believe that saving the ten has better consequences than saving the nine – after all, you save ten rather than nine! So consequentialism says you should save the ten. However, this is the wrong verdict. What's gone wrong here is clear: when ranking the option *save the ten*, consequentialism mistakenly ignores your doubts about being able to save the ten. That's because the outcome where you get to shaft A too late to save anyone doesn't count as a consequence of *save the ten*! The puzzle is to come up with an option set that makes consequentialism give the right verdict. (See Pollock 2006 and Hedden 2012.)

One response to the puzzle is what I call the *retreat strategy*. This approach retreats from overt action to the agent's inner life in order to find the options. On one variant, the options in the miners case are *intend to save the ten* and *intend to save the nine* (see Hedden 2012). The idea is that consequentialism ranks the latter best because you believe a possible consequence of intending to save the ten is travelling to shaft A only to get there too late to save anyone.

So on this conception of options, consequentialism says you should intend to save the nine – an acceptable verdict.

Alas, the retreat strategy fails. First, our theory of practical rationality should apply to agents whose actions aren't always preceded by intention. Second, the retreat strategy doesn't provide a general enough solution. Just as you can doubt your ability to save the ten, so you can doubt your ability to *intend to* save the ten – perhaps, for instance, the nine includes your family. Such doubts are relevant to the ranking of the intentions but consequentialism completely ignores them.

Where does this leave us? One option is to say that the sorts of cases causing trouble are “don't cares”. However, we have clear intuitions in such cases. And they aren't far-fetched. So we should expect our theory of practical rationality to apply to them. Another option is to look for an alternative way of ranking options (see Pollock 2006). But must we really abandon consequentialism because it's difficult to find suitable option sets?

Thankfully not. I propose an unusual conception of options to save consequentialism. The conception countenances making true a certain counterfactual as an option. In particular, in the miners case, save the ten is replaced by making true *if you could save the ten, then you would save the ten*. The other option is save the nine (or making true the corresponding counterfactual – it doesn't make a difference). The idea is that consequentialism ranks the counterfactual worse than *save the nine* because you believe one possible consequence of making it true is travelling to shaft A only to get there too late to save anyone. So consequentialism says you should save the nine – an acceptable verdict. Furthermore, this solution is not subject to the same worries as the retreat strategy: it need not assume intention always precedes action; and though it's possible to doubt that you can make true the counterfactual, this doubt is irrelevant to the counterfactual's ranking.

This solution leaves a few loose ends that I'll tie up. I'll mention two here. First: is a counterfactual the sort of thing that you can make true? I say yes, because you control the counterfactual's truth in virtue of controlling more ordinary facts such as the fact that you save the ten.

Second, my appeal to counterfactuals mirrors an appeal to counterfactuals by the moral luck sceptic. The sceptic says that agents are responsible in virtue of how they would act under various circumstances. However, a number of problems beset scepticism. For instance, the libertarian says that counterfactuals with free actions as consequents are guaranteed to be false. Do the same problems transfer across to my conception of options? I claim that the

analogous problems have easy solutions, in part because the antecedent-worlds for my counterfactuals are closer to the actual world than the antecedent-worlds for the sceptic's counterfactuals.

Works cited

Hedden, Brian. 2012. 'Options and the subjective ought', *Philosophical Studies*, 158: 343-60.

Pollock, John L. 2006. *Thinking about acting: Logical foundations for rational decision making* (Oxford University Press).