

## V Consequentialism

Preamble: What do normative theories describe?

Two very different models for normative ethics:

*Mathematics.* An a priori exploration of a domain that is robustly independent of us. But maths has a method of proof; and its success is tied in to natural science.

*Linguistics.* In some sense the subject matter is not independent of us. But there are still plenty of ways that we can go wrong (internal inconsistency; divergence between behaviour and belief about it; etc.)

A general approach here from response-dependent analyses: the special status of

X is F iff competent agents would react to X in a certain way

Compare secondary qualities like colour. What is the special status? A priori? Necessary?

### Consequentialism

Consequentialism is, very crudely, the view that what matters, when assessing the moral status of an action, are the consequences. Not just that the consequences are *one of the things* that matter (almost everyone thinks that consequences are important—if failing to divert the trolley means that millions of people will be killed, that is rather different to thinking that five will be killed; few agree with Kant about the axeman at the door); but that ultimately they are *all* that matters. ('Ultimately' since other things might matter instrumentally.)

General feature of consequentialist theories: the theory of the right is secondary to the theory of the good. That is, one judges whether an action is right by looking to the goodness of the state of affairs that it produces; the best action is the one that produces the best consequences. Or is it—a difficult issue—the one that one *expects*, or *reasonably expects*, to produce the best consequences? A distinction that one might draw here: evaluating the act, and evaluating the agent. Perhaps the latter should be done in terms of expected consequences.

Does everything collapse into consequentialism? Isn't 'doing the right thing' a kind of consequence? Williams: what matters is my doing the right thing (e.g. my keeping a promise) not the right thing being done (other people keeping promises). But can't that now be formulated as a consequence? (And doesn't it threaten making the opposition unacceptably agent-relative.) An unresolved issue: perhaps the best way to understand it is that the right must be *grounded* in the good—i.e. the duty must be grounded in the consequence—and this doesn't really show us how to do that. In any case, consequentialism is just a shell; it needs to be filled out if it is to give any policy recommendations. And the standard way of doing this is utilitarian.

## Utilitarianism

Utilitarianism can be seen as one version of consequentialism, characterized by two further features:

- (i) the relevant consequences are *human happiness, or human preference satisfaction*
- (ii) the requirement is that these consequences be *maximized*

Some general worries: the overdemandingness of consequentialism; the failure to accommodate partiality; the lack of space for supererogation.

But how significant are our reactions here? Mightn't they just show that we are biased in our judgments? Clean hands as moral self-indulgence. Josh Greene's fMRI studies—what, if anything, do they show? All this seems to make sense if mathematics is our model. But what if we take the linguistic model?

Going back to the two features of utilitarianism in turn: (i) what is so special about happiness?; happiness machines; what if it is not equally distributed—what if some people feel it more intensely than others? Alternatively: what is special about preferences?; what does addiction tell us about them? (ii) why maximize?; the repugnant conclusion; why not maximize the average? doesn't distribution matter? the separateness of persons.

## Possibility

But now a general defence: consequentialism is a doctrine that requires making things as good as we can. If human psychology entails that we *will in fact* behave in certain ways—we are partial etc—then we need to factor that in, and work around it.

Rule v Act utilitarianism. Overall we will do best by following certain rules of thumb. Cases of moral temptation. This involves getting things wrong at some times but not others.

More radically: utilitarianism as a self-effacing theory. Sidgwick. Doing the the right things for the wrong reasons. Government House mentality.

Perhaps none of us can accept utilitarianism at any time. This seems to give rise to a certain schism in our attitudes. But worse: now what reason do we have for believing in utilitarianism?