Prankster's Ethics

There's often good consequences to doing immoral acts. Well-timed bouts of immorality can entertain, provoke interesting discussion, and even inform. (Think of the artful prankster, or the amusing but offensive dinner guest.) These are all goods, and they may even be greater goods than the harm done by the immoral act. This suggests that what is right to do, and what is good to do, come apart. We distinguish the cases we have in mind from cases that are usually thought to be counterexamples to consequentialism. Careful analyses of the cases shows that also they reveal problems for theories that tie moral value too closely to moral character or to moral dispositions.

1. A Quick Argument for Boorishness

Diversity is a good thing. Some of its value is instrumental. Having people around with diverse beliefs, or customs, or tastes, can expand our horizons and potentially raise to salience some potential true beliefs, useful customs or apt tastes. Even diversity of error can be useful. Seeing other people fall away from the true and the useful in distinctive ways can immunise us against similar errors. And there are a variety of pleasant interactions, not least philosophical exchange, that wouldn't be possible unless some kinds of diversity existed. Diversity may also have intrinsic value. It may be that a society with diverse views, customs and tastes is simply

thereby a better society. But we will mostly focus on diversity's instrumental value here.

We think that what is true of these common types of diversity is also true of moral diversity. By moral diversity we mean not only diversity of moral views, though that is no doubt valuable, but diversity of moral behaviour. In a morally diverse society, at least some people will not conform as tightly to moral norms as others. In short, there will be some wrongdoers. To be sure, moral diversity has some costs, and too much of it is undoubtedly a bad thing. Having rapists and murderers adds to moral diversity (assuming, as we do, that most people are basically moral) but not in a way that is particularly valuable. Still, smaller amounts of moral diversity may be valuable, all things considered. It seems particularly clear that moral diversity within a subgroup has value, but sometimes society as a whole is better off for being morally diverse. Let us consider some examples.

Many violations of etiquette are not moral transgressions. Eating asparagus spears with one's fork is not sinful, just poor form. But more extreme violations may be sinful. Hurtful use of racial epithets, for example, is clearly immoral as well as a breach of etiquette. Even use of language that causes not hurt, but strong discomfort, may be morally wrong. Someone who uses an offensive term in polite company, say at a dinner party or in a professional philosophical forum, may be doing the wrong thing. (Let's name the term "Phil", although that may not be how it sounds.) But having the wrongdoer around may have valuable consequences. For example, they generate stories that can be told, to great amusement, at subsequent dinner parties. They also prompt us to reconsider the basis for the standards we ourselves adopt in such matters. The reconsideration may cause us to abandon useless practices, and it may reinforce useful practices. These benefits seem to outweigh the disutility of the discomfort felt by those in attendance when the fateful word drops from the speaker's lips. These side benefits do not make the original action morally permissible. Indeed, it is precisely because the action is not morally permissible that the benefits accrue.

While we think that case is one of valuable moral diversity, some may question the *immorality* of the act in question. So let us try a more clearly immoral case: the mostly harmless prankster. Sam is a pie-thrower. Sam doesn't just throw pies at the rich and infamous. No, Sam's pies land on common folk like you and I, often for no reason beyond Sam's amusement. Causing gratuitous harm for one's own amusement is immoral. And a pie in the face, while better than a poke in the eye with a burnt stick, is harmful. But it may, in some circumstances, have side benefits. There will be the (guilty) pleasure occasioned in the unharmed bystanders, though it would be wrong to put too much weight on that. Other more significant benefits may accrue if Sam's society is otherwise saintly. Sam's existence will prompt people to take some simple, and worthwhile, precautions against perpetrators of such attacks. Even if society currently contains no malfeasants, such precautions will be useful against future wrongdoers. This benefit will increase if Sam graduates from pie-throwing to more varied pranks. (As may the entertainment value of Sam's pranks.) Many computer hackers perform just this function in the present world. Malicious hackers on the whole cause more harm than good. But other hackers, who hack without gratuitously harming, provide a protective benefit by informing us of our weaknesses. These are the pie-throwers of the virtual world. Sam's actions have other benefits. If Sam's pranks are harmless enough, some will mistakenly think that they are morally acceptable, and we can have enjoyable, valuable, philosophical discussions with them. (Note that this benefit also increases if Sam varies the pranks.) The upshot is that Sam's pranks can make the world a better place, all things considered, despite being immoral. Indeed, in some ways they make the world a better place *because* they are immoral.

The philosophical point, or points, here may be familiar. One point certainly *is* familiar: we have here an example of a Moorean organic unity. The goodness of the whole is no simple function of the goodness of the parts. It might be thought that follows simply from the familiar counterexamples to utilitarianism, and that our

examples have no more philosophical interest than those old counterexamples. Both of these thoughts would be mistaken.

The familiar counterexamples we have in mind include, for example, the case of the doctor who kills a healthy patient to harvest her organs, or the judge who executes an innocent man to prevent a riot. Importantly, those examples do not refute consequentialism in general, but only a version of consequentialism that adopts a particular kind of reductive analysis of the good. The details of the analysis won't matter here, but it may be an analysis of goodness in terms on happiness, or preference satisfaction. If we give up the reductive analysis of goodness, we can say that the doctor and the judge do not make for a better society. A familiar heuristic supports that claim. (We take no stand here on whether this heuristic can be turned into an analysis.) Behind the Rawlsian veil of ignorance, we would prefer that there not be such doctors or judges in society. We think that most of us would agree, even in full appreciation of the possibility that we will be saved by the doctor, or possibly the judge. On the other hand, we think we'd prefer a society with the occasional boorish dinner guest, or a rare pie-thrower, to a society of moral saints. We say this in full appreciation of the possibility that we may get a pie in the face for our troubles. Possibly if we *knew* we would be the pie-throwee we would change our minds, but fortunately pies cannot penetrate the veil of ignorance.

So a consequentialist can avoid the utilitarian counterexamples simply by saying less about goodness. The most natural retreat is to behind the veil of ignorance, but our examples can reach even there. This is far from the only interesting consequence of the examples.

2. The Good, the Right, and the Saintly

We think that the cases the curser and the pie-thrower are examples of situations in which (a) an agent ought not to φ, and (b) it's best that the agent *does* φ. Our judgements about the cases are not based on any theoretical analysis of the right and the good. They're simply intuitions about cases—it just seems to us that the right

thing to say about the pie thrower is that she ought not to do what she does, but that it's still best if she does it. To the extent that these intuitions are puzzling or theoretically problematic (and we think that they are at least a little bit puzzling, and at least potentially problematic), it's open to us to reject one or the other intuition about the cases, and either deny that the curser and the pie thrower ought not to curse or throw pies, or deny that it's best that they *do* curse and throw pies. This is an option, but we think it's not a very attractive one. Suppose that instead we take the intuitions at face value, and accept our judgements about the cases. What follows?

Our analysis of the examples is incompatible with two attractive views about the connection between *goodness* (that is, the property of things—in particular worlds—in virtue of which some of them stand in the *better than* relation to others) and *rightness*, and between goodness and good character:

1) It's better if everyone does what's right.

2) It's better if everyone has good character.[1]

Now, neither of these will do as a philosophical thesis. But it's probably not worth spending the time and effort on patching them up, since even the patched-up versions will be false.

If the pie-thrower ought not to throw her pies, but nonetheless it's best that she does, no patched-up version of (1) that captures the intuition behind it can be right. Any patched-up version of (1) will still be claiming that there's a very tight connection between what it would be *right* for us to do (what we *ought* to do) and what it would be *best* for us to do. Any plausible elaboration on (1) will include a commitment to the thesis that, if we ought not to do something, then it's best if we

[1] (2) is quite a natural position to hold if one is trying to capture the insights of virtue ethics in a consequentialist framework, as in Driver (2001) or Hurka (2001). But if we take 'better' in a more neutral way, so (2) does not mean that there are better consequences if everyone has good character, but simply that the world is a better place if this is so, even if this has few consequences, or even negative consequences, then it will be a position common to most virtue ethicists.

don't do it. But if our analyses of the cases of the dinner guest and the pie-thrower are right, then these are counterexamples.

What about (2)? Well, it's *not* better if the cursing dinner guest has good character. What happens if we suppose that the guest *does* have good character? One of two things: (i) He'll no longer curse at dinner parties, and we'll lose the benefits that come from his cursing. This would be bad. (ii) He'll still curse at dinner parties, but he'll be cursing in a *studied* way. He'll be cursing because he's seen that things will be better if *somebody* uses foul language in inappropriate circumstances, and he's taken it upon himself to fill the unfilled functional role. This would also be bad. This sort of *studied* bad conduct doesn't have the same value as bad conduct that springs from bad character. Here is some evidence for this: We value the curser's breaching of societal norms, even though he ought not to do it. Were we to *find out* that every expletive had been *studied*, produced either to produce these important social goods, or to create a familiar bad-boy image, we would *stop* valuing his breachings of the moral order. They would, instead, become merely tiresome and annoying. Since we value spontaneous "Phil"-shoutings which are products of less-than-optimal character, but we do *not* value studied "Phil"-shoutings which are products of exemplary character, it's very plausible to conclude (though admittedly not quite mandatory) that the spontaneous curses are much more valuable than the studied ones. We're inclined to say, in fact, that while having a few *spontaneous* cursers around makes things better, having *studied* cursers around makes things worse. Since you have to have less-than-perfect character in order to be a spontaneous curser, it follows that you can't get the benefits of having cursers around without having some people with less-than-perfect character around. And since it's better to have the cursers than not, it's better to have some people with less-than-perfect character around than not. This will be incompatible with almost any plausible way of cashing out (2).[2]

[2] Specifically, it will be incompatible with any *maximizing* version of (2). There might be 'threshold' versions of (2) that don't fall afoul of this kind of problem because they don't claim it would be best for everyone to have *perfect* character, but only that it would be best for everyone to have *pretty good* character, or at least for nobody to have *really bad* character.

3. A Problem about Quantifier Scope?

But isn't there a sense in which (for example) the pie-thrower *ought* to throw his pies? After all, if nobody was throwing pies, we might think to ourselves, "gosh, it would be better if there were a few—not many, but a few—pie throwers around". Then it would be natural to conclude, "somebody ought to start throwing pies at strangers". And then it would be natural to infer that at least the first person to start throwing pies at strangers would be doing what they ought. It would be natural, but it would be wrong. The plausible reading of "someone ought to start throwing pies at strangers" is, "it ought to be that somebody starts throwing pies at strangers", not, "there's somebody out there such that they ought to start throwing pies at strangers". So we haven't gotten anybody a moral license to throw pies yet. And in fact it's very plausible that the right way to understand claims of the form, ⌜It ought to be that P⌝ is as meaning that it would be *better* if it were the case that P; that is, as making claims about what would be *good*, not about what would be *right*.

There's a puzzle about what to make of cases where we're inclined to say that it ought to be that somebody φs—that is, that somebody ought to φ; but also that there's nobody such that *they* ought to φ—in fact, that everybody is such that they ought *not* to φ. [3] Maybe the fact that our intuitions about the examples give rise to these kinds of puzzling cases is evidence that one or the other of our intuitions ought to be rejected. The move we suggested above is that the reason this seems so puzzling is that we've been punning on "ought". The "ought" in "somebody ought to start throwing pies" doesn't have anything much to do with what moral obligations anybody has—doesn't have anything much to do with what's *right*—but has a great deal to do with what's *good*. And if that's the case, then all we have is more evidence against the tight connection between the right and the good: it would be better if

[3] It's actually the second part that makes it puzzling. Compare the familiar and unproblematic situation in which we ought to give you a horse, but there's no horse such that we ought to give you *that* one, and the more troubling situation in which we ought to give you a horse, but every horse is such that we ought *not* to give you that one.

somebody started throwing pies, but everybody has a moral obligation not to. So it would be better if somebody did what they oughtn't.

4. Value, Desire and Advice

Although this "ought" has little to do with what's *right*, it might have a lot to do with what we find desirable. And this will cause problems for some familiar meta-ethical theories. Quite naturally, Jack does not desire to throw pies at strangers for amusement in the actual world. Jack's a very civic minded fellow in that respect. In fact, his concern for others goes deeper than that. He'd be quite prepared to risk his body for the sake of his fellow citizens. As it turns out, he's been a volunteer fire fighter for years now. And Jack likes to think that if need be, he would be prepared, to use an old fashioned phrase, to risk his soul for the community. He hopes he would be morally depraved if what the society needed was depravity. Jack agrees with the discussion of character in section 2, so he hopes that when society needs a pie-thrower, he will step up with the plate, and do so directly because he wants to throw pies at innocent bystanders. Letting C stand for the circumstances described above, where it would be good for there to be more wrongdoing, Jack's position can be summarised by saying that he desires that in C he desires that *he* throws pies at innocents.

Does this all mean Jack values his throwing pies at innocents in C? Not necessarily. Does it mean that if we were all like Jack, and we are subjectivists about what is right, it would be right to throw pies at innocents in C? Definitely not. David Lewis (1989) equates what we value with what we desire to desire.[4] And he equates what is valuable with what we value. The text is not transparent, but it seems Lewis wants *valuable* to subsume both what we call the 'right' and the 'good'. And this he cannot have. Assume that everyone in Jack's community desires to (*de se*) desire that

[4] More precisely, with what we desire to desire in circumstances of appropriate imaginative acquaintance. We can suppose that Jack, and everyone else under discussion in this paragraph, is suitably imaginatively acquainted with the salient situations. Jack knows full well what it is like to get a pie in the face.

(s)he throw pies at innocents in C. That does not make it right that pies are thrown at innocents. We take no stand here on whether the flaw is in the equation of personal value with second-order desire, or in the reduction of both rightness and goodness to personal value, but there is a problem for Lewis's dispositional theory of value.[5]

This point generalises to cause difficulties for several dispositional theories of value. For example, Michael Smith (1994) holds that right actions are what our perfectly rational selves would advise us to do. This assumes that when the good and the right come apart, our perfectly rational selves would choose the right over the good. And it's far from clear that Smith has the resources to argue for this assumption. Smith's argument that our perfectly rational selves will advise us to do what is right relies on his earlier argument that anyone who does not do what she judges to be *right* is practically irrational, unlike presumably our perfectly rational selves. And the main argument for that principle is that it is the best explanation of why actually good people are motivated to do what they judge to be right, even when they change their judgements about what is right. But now we should be able to see that there's an alternative explanation available. Actually good people might be motivated to do what they judge to be *good* rather than *right*. We have seen no reason to believe that the right and the good actually come *radically* apart, so this is just as good an explanation of the behaviour actual moral agents as Smith's explanation. So for all Smith has argued, one might judge φing to be right, also judge it not to be good, hence be not motivated to φ, and not be practically irrational. Indeed, our

[5] Someone might think it obvious that Lewisian value can't be used in an analysis of both rightness and goodness, since it is *one* concept and we are analysing *two* concepts. But Lewisian value bifurcates in a way that one might think it is suitable for analysing both rightness and goodness. Since there are both *de dicto* and *de se* desires, one can easily draw out both *de dicto* and *de se* values. And it is *prima facie* plausible that the *de dicto* values correspond to what is good, and the *de se* values to what is right. Indeed, given a weak version of consequentialism where these two can be guaranteed to not directly conflict, this correspondence may well hold. But we think the pie-thrower threatens even those consequentialists. The net philosophical conclusion is that the pie-thrower is a problem for Lewis's meta-ethics, but only because (a) she is a problem for Lewis's consequentialism, and, surprisingly, (b) Lewis's meta-ethics depends on his consequentialism being at least roughly right.

perfectly rational self might be just like this.[6] Hence we cannot rely on our perfectly rational self to be a barometer of what is right, as opposed to what is good.

References

Driver, Julia (2001) *Uneasy Virtues*. Cambridge: Cambridge University Press.

Hurka, Thomas (2001) "Vices as Higher-Level Evils" *Utilitas* 13: 195-212.

Lewis, David (1989) "Dispositional Theories of Value" *Proceedings of the Aristotelian Society* Supplementary Volume 63: 113-37.

Smith, Michael (1994) *The Moral Problem*. Oxford: Blackwell.

[6] We have glossed over a technical point here that is irrelevant to the current discussion. What matters is not whether our perfectly rational selves are motivated to φ, it matters whether they desire that we φ, and hence whether they are motivated to advise us to φ. Keeping this point clear matters for all sorts of purposes, but not we think the present one.