

# Immodest Consequentialism and Character

MICHAEL SMITH

*Australian National University*

The fact that we place the value that we do on the traits of character constitutive of being a good friend, and the acts that good friends are disposed to perform, creates a considerable problem for what I call 'immodest global consequentialism'. The problem is, in essence, that the very best that the immodest global consequentialists can do by way of vindicating our most deeply held convictions about the value of these traits of character and actions isn't good enough, because, while vindicating our possession of those convictions, the attempted vindication undermines the truth of the convictions thus possessed. This is especially bad news because, as I argue, immodest global consequentialism is the only version of consequentialism that can be distinguished in any principled way from a form of non-consequentialism.

A plausible normative theory would underwrite our most deeply held normative convictions. So much is uncontroversial. However, given that we have deeply held convictions about the value of various traits of character and the actions that those traits of character occasion, many think that this uncontroversial thesis in moral epistemology provides the basis of a powerful line of argument against consequentialism. The charge, more specifically, is that consequentialism is implausible because it is inconsistent with our most deeply held convictions about the value of the character traits constitutive of being a good friend and the actions that those who are good friends are disposed to perform.<sup>1</sup>

My main aim in the present paper is to evaluate this charge against consequentialism. To anticipate, my argument will be that once we distinguish between two kinds of consequentialism – a modest and an immodest version – we see that though the fact that we place the value we do on the traits of character constitutive of being a good friend creates no problem at all for the modest theory, it creates a considerable problem for the theory that is more immodest. This might initially sound like good news. Perhaps consequentialists should go in for a little modesty. But it is in fact bad news. For, I argue, there is no principled difference between being a modest consequentialist and being a non-consequentialist. The fact that we place the value we do on the traits of character constitutive of being a good friend and the actions that good friends are disposed to perform thus creates a considerable

<sup>1</sup> Michael Stocker, 'The Schizophrenia of Modern Ethical Theories', *Journal of Philosophy*, lxxiii (1976); Dean Cocking and Justin Oakley, 'Indirect Consequentialism, Friendship, and the Problem of Alienation', *Ethics*, cvi (1995); Thomas Scanlon, *What We Owe to Each Other*, Cambridge, MA, 1998, pp. 88–90.

problem for the only version of consequentialism deserving of the name, that is, for immodest consequentialism.

As is perhaps clear, I thus have a secondary aim in the present paper as well, and that is to put immodest consequentialism firmly on the table for subsequent discussion. Immodest consequentialism is, as I see things, the version of consequentialism that those with consequentialist leanings should be in the business of defending. Yet immodest consequentialism is all but ignored in the literature.<sup>2</sup> My hope is that the present paper will help focus the attention and efforts of consequentialists on a defence of the immodest version of their theory.

## I. THE NATURE OF CONSEQUENTIALISM

Many different theories can be grouped under the label 'consequentialism'. Let me therefore begin by spelling out more precisely the nature of the theory that is under attack.

As I understand it, consequentialism is, at bottom, a theory of moral evaluation. It says that the morally best *x*, for some *x* in the category of evaluand, is the *x* that has the morally best consequences. Moral value might be intrinsic to the evaluand in question, the consequence being the obtaining of the evaluand itself; or it might be extrinsic, accruing to the evaluand in virtue of the causal (or some other) relations in which the evaluand stands to other states or events. Consequentialism itself takes no stand on this. Similarly, certain moral values might depend on what the subject involved expects the consequences of the evaluand to be; or it might depend on what the subject involved would expect the consequences to be, if his expectations had been formed appropriately in the light of the available evidence; or it might depend on what the actual consequences are. Again, consequentialism itself takes no stand on this.

Crucially for what follows, however, consequentialism is not silent about the nature of moral value. A distinctive mark of a consequentialist theory, as opposed to a non-consequentialist theory, is that it takes moral values to be neutral, as opposed to relative. Though this distinction is probably familiar enough, an example will help fix more precisely the idea that I have in mind. Consider pleasure. Pleasure can be either a neutral value or a relative value. When we take pleasure to be a neutral value we think that the world is a better place when pleasure is experienced in it, independently of who experiences that pleasure. When we take pleasure to be a relative value, by contrast, we

<sup>2</sup> The exception is Frank Jackson, 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics* (1991).

deny this to be so. We think that it is one's experiencing pleasure oneself that makes the world a better place, or perhaps that it is one's friends' experiencing it, or a member of one's family.

More precisely, pleasure is a neutral value when the value of a possible world, as judged in terms of pleasure, can be identified without knowing which of the subjects experiencing pleasure in that possible world one is oneself, or which of those in the world are those to whom one bears some special relationship. Pleasure is a relative value when the value of a possible world cannot be thus identified. One would first need to know how much pleasure one is experiencing oneself, or perhaps how much one's friends or family members are experiencing, before one could judge the value of a world.

Generalizing, neutral values are those that permit us to rank possible worlds in terms of their value without our having any knowledge of the relationship which the objects in those possible worlds stand in to ourselves. Relative values, by contrast, are those that do not permit such a ranking. Rankings of possible worlds in terms of relative values require such indexical knowledge. A distinctive mark of a consequentialist theory is thus that it says that the morally best  $x$ , for some  $x$  in the category of evaluand, is the  $x$  that has the morally best consequences, where the best consequences in turn are those that maximize neutral values.

I have said that consequentialism is a theory which says that the morally best  $x$ , for *some*  $x$  in the category of evaluand, is the  $x$  which has the morally best consequences. There is, however, a major divide within consequentialist theories. This is the divide between those theories which Philip Pettit and I have elsewhere called *local*, as opposed to those that are *global*.<sup>3</sup>

Global consequentialism, as the name implies, is a global theory of moral evaluation. It identifies the morally best  $x$ , for *any*  $x$  in the category of evaluand, as that  $x$  which maximizes moral value. Local consequentialism, by contrast, privileges *some*  $x$  in the category of evaluand, evaluating it in the same way as global consequentialism, but then goes on to evaluate all other evaluands in terms of some relationship in which they stand to that  $x$ , that is, to the evaluand in the privileged category. Some examples will illustrate.

According to global consequentialism, the best *act* for someone to perform is the act that has greater moral value than any of the acts that might have been performed instead. This, in turn, is the act which is morally right. Moreover, according to global consequentialism, in

<sup>3</sup> Philip Pettit and Michael Smith, 'Global Consequentialism', *Morality, Rules, and Consequences: A Critical Reader*, ed. Brad Hooker, Elinor Mason and Dale E. Miller, Edinburgh, 2000, pp. 121–33.

order to figure out what the morally best *motive-set* for someone to possess is we must simply reapply the consequentialist test, this time to the possession of motives. Instead of asking whether *the act performed* is better than any of the acts that might have been performed instead, we ask whether *the set of motives possessed* is better than any of the sets of motives that might have been possessed instead. And when it comes to evaluating *traits of character*, global consequentialism tells us once again to reapply the consequentialist test: the morally best set of character traits for someone to possess is that set of character traits the possession of which has greater agent-neutral moral value than any of the sets of character traits that they might have possessed instead.

Global consequentialism thus contrasts with all forms of consequentialism that privilege one or another category of evaluand. This is the distinctive feature of local consequentialism. For example, local *motive* consequentialism privileges the category of motives, defining the morally best motives as those whose possession maximizes neutral value. In this it agrees with global consequentialism. But it then goes on to stipulate that the morally best acts, the morally best set of character traits, and so on, are those that bear some appropriate relation to the morally best motives. For example, that relation might be such that local *motive* consequentialism tells us that the morally best act to perform is an act produced by the morally best motive.<sup>4</sup>

The contrast between the global consequentialist evaluation of acts the local *motive* consequentialist evaluation should be plain. Global consequentialism holds that the moral value of acting on a desire whose possession maximizes neutral value must be determined empirically, just like the moral value of anything else. The issue is whether acting on a desire whose possession maximizes neutral value does indeed maximize neutral value. Local *motive* consequentialism, on the other hand, insists that only the moral value of motives needs to be determined empirically, the moral value of everything else following by stipulation.

Another example. Local *act* consequentialism privileges the category of acts, defining the morally best acts as those which maximize neutral value. In this it agrees with global consequentialism. But it then goes on to identify the morally best motive-set, the morally best

<sup>4</sup> Though why shouldn't the local *motive* consequentialists say instead that the relation is one according to which the morally best acts are those that would have been performed by people if they had had the morally best motives? Or why not say instead that the relation is one according to which the morally best acts are those that would have been performed by people if they had had the set of motives that it would have been morally best for them to try to acquire? It is because we see no principled and plausible way of answering these questions that Pettit and I reject local *motive* consequentialism (Pettit and Smith).

set of character traits, and so on, as those sets that bear some appropriate relation to the morally best acts. For example, that relation might be such that local *act* consequentialism tells us that the morally best motive set is that set which motivates people to perform all and only the morally best acts.<sup>5</sup> This is presumably a set comprising a single motive: the desire to act so as to maximize neutral value. Moreover local *act* consequentialism similarly defines the morally best set of character traits. On certain assumptions about the nature of the neutral moral values this would mean that, for example, benevolence is the morally best trait of character someone could possess.<sup>6</sup>

The contrast between the global consequentialist evaluation of motives and character traits and the local *act* consequentialist evaluation should be plain. Global consequentialism holds that the moral value of possessing the desire to maximize neutral value, or that of being benevolent, must be determined empirically, just like the moral value of anything else. The issue, in each case, is whether the possession of the desire to maximize neutral value, or the character trait of benevolence, maximizes neutral value. Local *act* consequentialism, on the other hand, insists that only the moral value of acts needs to be determined empirically, the moral value of everything else following by stipulation.

It will perhaps come as no surprise to hear that local consequentialism, in all its varieties, is plagued with problems. The main problem, in essence, is that the local consequentialist needs to give some principled reason for assessing the moral value of some evaluands in one way (namely, via a direct application of the consequentialist principle) and other evaluands in a completely different way (namely, via a stipulation given a certain sort of relation). However, since Philip Pettit and I have argued this point at length elsewhere,<sup>7</sup> and since global consequentialism is in any case the version of consequentialism that some of the best known consequentialists defend,<sup>8</sup> I will not take

<sup>5</sup> Local act consequentialists might defend the choice of this relation by citing the following remark of Thomas Nagel's. '[A]n impersonal morality requires of us not only certain forms of conduct but also the motives required to produce that conduct. This much is I think true of any morality properly worked out. If we are required to do certain things, then we are required to be the kinds of people who will do those things' (Nagel, *The View from Nowhere*, Oxford, 1986, p. 191). Needless to say, the global consequentialist denies that this is true. Rather, according to the global consequentialist, all requirements are generated by the consequentialist principle itself

<sup>6</sup> Compare Philippa Foot, 'Utilitarianism and the Virtues', repr. *Consequentialism and its Critics*, ed. Samuel Scheffler, Oxford, 1988, pp. 224–42.

<sup>7</sup> Pettit and Smith.

<sup>8</sup> Henry Sidgwick, *The Methods of Ethics*, 7th edn. London, 1907; Derek Parfit, *Reasons and Persons*, Oxford, 1984; Peter Railton, 'Alienation, Consequentialism, and the Demands of Morality', repr. *Consequentialism and its Critics*, ed. Samuel Scheffler, Oxford, 1988; 'How Thinking about Character and Utilitarianism Might Lead to Re-

the time to repeat those arguments here. In what follows I will simply take it for granted that the most plausible form of consequentialism is global consequentialism.

## II. A FIRST ATTEMPT AT STATING THE CHARGE AGAINST CONSEQUENTIALISM

The focus, in the remainder, is to be on global consequentialism's evaluation of various traits of character and the actions that those traits of character occasion. The question to be addressed, more specifically, is whether global consequentialism can adequately account for our most deeply held convictions about the value of the traits of character constitutive of being a good friend and the actions that good friends are disposed to perform. Before addressing this question, however, we must first ask what these traits of character are supposed to be.

There are, it seems to me, two related features of the character traits constitutive of being a good friend on which we need to focus. The first is that, when faced with a choice between giving a benefit to her friend, or a slightly greater benefit to a complete stranger, then, assuming that the consequences of her actions are otherwise identical, a good friend will choose to give the benefit to her friend. Since this is a claim about the relative strength of the motivations of someone who is a good friend let's call it the 'Friendly Motivation Thesis'. The second feature is that a good friend is disposed to offer the fact that her friend is *her friend* as a justifying reason for her doing what she is motivated to do, given the truth of the Friendly Motivation Thesis. This differs from the Friendly Motivation Thesis in being a claim not about the relative strength of the motivations of a good friend, but about the relative weight she assigns to the value of benefiting her friends as opposed to complete strangers. Accordingly, let's call this the 'Friendly Justification Thesis'.

With these two features of the traits of character constitutive of being a good friend clearly in view we are in a position to ask the crucial question. Does global consequentialism accord with the value we ordinarily place on the character traits constitutive of being a good friend and the actions that good friends are disposed to perform?

In what follows I propose to assume, much as others do, that we ordinarily place great value on the two dispositions just described in the Friendly Motivation Thesis and the Friendly Justification Thesis. Moreover I will also assume that the reason that we ordinarily place

the value that we do on these two dispositions is that we take the dispositions to reflect the values that are at stake: we take it to be good to prefer providing the benefits that we do to our friends, and to offer the fact that our friends are our friends as a justifying reason for this preference, because the fact that our friends are our friends *is* a justifying reason both for the provision of such benefits and for being so motivated.

Let's now ask what global consequentialism has to say about the value of the two dispositions and of the acts that these dispositions occasion. The answer, unsurprisingly, is that global consequentialism holds that the moral value of these two dispositions and the acts they occasion are to be determined empirically. Since I am in no position to argue about the empirical facts, in what follows I will simply assume, much as those in the literature do, that our actual empirical circumstances are such that, when we apply the global consequentialist test to these two traits of character, it turns out that their possession maximizes neutral value.<sup>9</sup> Moreover I will also simply assume that on many occasions on which we act on these dispositions, we fail to maximize neutral value.

It might be thought that once we make this empirical concession it follows that, far from there being any great tension between the global consequentialist's assessment of the value of the traits of character constitutive of being a good friend and the acts that good friends perform, on the one hand, and our most deeply held convictions, on the other, there is in fact a good deal of convergence. True enough, global consequentialism does tell us that the acts that good friends are disposed to perform will, on occasion, be wrong. But since it also suggests that it is morally good for them to be so disposed, and that it is morally good for them to believe what they believe about the justifying reasons they have for being so disposed, where, exactly, is the great tension supposed to be? For this accords exactly with our most deeply held convictions.

The answer emerges when we look more closely at what global consequentialism deems to be morally best. Given the Friendly Justification Thesis, global consequentialism tells us that it is morally best for someone to be disposed to offer the fact that her friend is her friend as a justifying reason for her giving a smaller benefit to her friend as

<sup>9</sup> Though I make this assumption I must say that it does seem to me to be rather unrealistic. In the context of human misery on the scale that currently exists in the world it would, I think, be far more realistic to suppose that it would maximize neutral value for many of those of us who have a large number of resources at our disposal to be motivated almost exclusively to bring about as much neutral value as possible. A good number of us, at any rate, should remain friendless for the sake of the greater good. But I propose to leave this objection to one side for the purposes of the present argument.

opposed to a greater benefit to a stranger. But on the plausible assumption that a justifying reason for acting in one way rather than another is constituted by the fact that it would be better to act in the one way rather than the other<sup>10</sup> it would therefore seem to follow that someone who is so disposed is someone who assigns *relative* value to the welfare of her friends. For in ranking possible worlds in terms of their value she would need to have indexical knowledge: she would need to know who *her friends* are in the possible worlds she is ranking and who, in those worlds, the complete strangers are.

Indeed, not only is someone who is disposed to offer the fact that her friend is her friend as a justifying reason for giving a slightly smaller benefit to her friend, as opposed to a slightly greater benefit to a complete stranger, someone who assigns relative value to the welfare of her friends, she is someone who in addition thinks that such a gain in relative value of the welfare of her friends is of greater evaluative significance, *simpliciter*, than the loss in neutral value of providing for the welfare of the complete stranger.<sup>11</sup> The justifying reason in question is, after all, a reason for *preferring* giving the slightly smaller benefit to her friend to giving the slightly larger benefit to a complete stranger. But since, according to global consequentialism, moral values are one and all neutral, it would thus seem to follow that the theory is committed to its being morally best for us to have *false* beliefs about the relative value, *simpliciter*, of giving a slightly smaller benefit to our friends, as opposed to a slightly greater benefit to a complete stranger. After all, how could global consequentialism be consistent with its being better, *simpliciter*, for us to give a smaller benefit to a friend than a larger benefit to a complete stranger? The global consequentialist holds that it is morally best for us to provide the larger benefit to the complete stranger.

Here, then, lies the tension between global consequentialism's evaluation of the traits of character constitutive of being a good friend and our most deeply held convictions about the value of such traits. While it can be agreed that both global consequentialism and common sense deem it to be good to possess the traits of character constitutive of being a good friend, it is no part of common sense to hold that the beliefs described in the Friendly Justification Thesis are false. Indeed, as we have seen, quite the opposite is the case. We ordinarily hold that the reason it is morally good to have the beliefs described in the

<sup>10</sup> Michael Smith, *The Moral Problem*, Oxford, 1994, p. 95 and ch. 5.

<sup>11</sup> What does talk of evaluative significance, *simpliciter*, amount to? I take it that the intuitive idea is plain enough: it is the ranking of possible worlds that results when an evaluator who has one ranking of possible worlds in neutral terms and another ranking of possible worlds in relative terms merges those two rankings into one overall ranking. A more precise suggestion is made in note 27 below.

Friendly Justification Thesis, beliefs about the relative value of giving benefits to our friends, is that those beliefs are true: it is morally good to be disposed to offer the fact that your friends are your friends as a justifying reason for giving a slightly smaller benefit to them as opposed to a slightly greater benefit to a complete stranger, in cases where the consequences of your actions are otherwise identical, because the welfare of your friends does in fact have relative value, and, given the magnitude of the relative value realized by so acting, this option is indeed of greater value, *simpliciter*, than that of giving the larger benefit to a complete stranger. According to global consequentialism, however, none of this is true. The justification is rather that falsely believing that there is such relative value maximizes neutral value. The mismatch with our most deeply held convictions couldn't be more apparent.

### III. MODEST VERSUS IMMODEST GLOBAL CONSEQUENTIALISM

The global consequentialist has an obvious line of reply to the charge, as just stated. In getting clear about this reply we will, however, come to a much better appreciation of the real charge against global consequentialism.

The global consequentialist will begin with a reminder. Global consequentialism's account of the rightness of acts follows from two distinct claims. The first is the claim that an act is morally right if and only if it maximizes *moral* value. The second is the claim that all *moral* value is neutral. It thus follows that global consequentialism, as such, is simply silent as to whether or not there are any *non-moral* relative values. All that global consequentialism entails is that, if there are any non-moral relative values then, given that these values are *non-moral*, they are irrelevant to the determination of which acts are right and which are wrong.

In order to see what the distinction between moral and non-moral values amounts to, at least as the global consequentialist sees things, and to appreciate the force of his making this distinction in the context of the charge as just stated against his theory, it will be helpful if we consider once again the example of pleasure. As we have already seen, pleasure can be either a neutral value or a relative value. Pleasure is a neutral value when the experience of pleasure, no matter who experiences it, is deemed to make the world a better place. It is a relative value when the experience of pleasure by someone suitably related to oneself – for example pleasure experienced by oneself, or pleasure experienced by a member of one's family, or pleasure experienced by a friend – is deemed to make the world a better place.

But, of course, we needn't suppose that this is an either/or situation. Pleasure might be *both* a neutral *and* a relative value. Indeed, we would ordinarily suppose that it is both, as we ordinarily take ourselves to have at least two reasons for pursuing our own pleasure. One reason is that our own pleasure contributes to the total amount of pleasure in the world, independently of who experiences it: this reason is available to anyone. The other reason is that our own pleasure contributes to the sum total of pleasure that we ourselves experience: this reason is essentially available only to ourselves. This is why, when pleasure is the only value at stake in some choice situation, we feel quite justified in deciding between alternatives that contribute equally to the amount of pleasure in the world on the basis of which alternative contributes most to our own pleasure.

Global consequentialists will insist that this distinction is crucially important in the present context because it means that when their theory tells us that it is morally best for people to assign relative value to the welfare of their friends, it mustn't be assumed that it thereby tells us that it is morally best for people to assign a form of *moral* value to the welfare of their friends. It can instead be supposed to tell us that it is morally best for us to assign *non-moral* value to the welfare of our friends. If this is what his theory says then the global consequentialist can conclude that, notwithstanding the fact that global consequentialism holds that all moral value is neutral, this assignment of non-moral relative value to the welfare of our friends might even be true.

Moreover, armed with this distinction between moral and non-moral value, the global consequentialist is even in a position to say that it is consistent with his theory to hold that there is greater evaluative significance, *simpliciter*, in giving a slightly smaller benefit to our friends than there is in giving a slightly larger benefit to a complete stranger. The example of pleasure is once again illustrative. For once it is agreed that our own pleasure has non-moral relative value then, absent an argument to the effect that the neutral moral value of pleasure is lexically prior to the relative non-moral value of pleasure, the global consequentialist is bound to admit the possibility of cases in which we have to choose between maximizing pleasure in the sense of maximizing the neutral value associated with pleasure, independently of who experiences it, on the one hand, and maximizing pleasure in the sense of maximizing the relative non-moral value associated with the pleasure that we ourselves experience, on the other. There will therefore be at least possible cases in which the gain in relative non-moral value associated with the gain of pleasure that we ourselves experience is of greater evaluative significance, *simpliciter*, than the loss of neutral moral value associated with the loss of pleasure independently of who experiences it. We need simply to imagine a case in which the

gain, in relative terms, of the relative non-moral value of pleasure is sufficiently large, and the loss, in neutral terms, of the neutral moral value of pleasure is sufficiently small.

The upshot is thus that there are supplementary claims a global consequentialist can make which will allow him to embrace the Friendly Justification Thesis much as we ordinarily understand it. The first is the claim that there are, in addition to neutral moral values, relative non-moral values. These are the values at stake when we justifiably give a smaller benefit to a friend rather than a larger benefit to a complete stranger. The second is that such relative non-moral values as are at stake in a particular choice situation, like the situation of choosing between friends and strangers, can sometimes have so much evaluative significance, *simpliciter*, that they will outweigh the neutral moral values at stake in that choice situation. In order to distinguish the global consequentialist who accepts these two claims from those who might reject them let me call such a theorist, for obvious reasons, a *modest global consequentialist*. He is modest because he acknowledges the existence and significance of relative non-moral values.

According to the modest global consequentialist there are therefore at least two things to say about the value of the belief described in the Friendly Justification Thesis – that is, the belief that the fact that our friends are our friends is a justifying reason for preferring giving them a slightly smaller benefit to giving a slightly greater benefit to a complete stranger. The first thing to say is that it is morally good to possess this belief, morally good because its possession maximizes neutral value. The second thing to say is that the belief is true. The belief is true because, according to the modest global consequentialist, there is indeed relative non-moral value in the provision of smaller benefits to friends over larger benefits to complete strangers, and because, in contexts in which the consequences of our actions are otherwise identical, there is indeed more evaluative significance, *simpliciter*, involved in giving a slightly smaller benefit to a friend, an action which has relative non-moral value, than there is in giving a slightly larger benefit to a complete stranger, an action which has neutral moral value. The modest global consequentialist's view is thus consistent with the Friendly Justification Thesis, much as we ordinarily understand it.

#### IV. THE PROBLEM WITH MODEST GLOBAL CONSEQUENTIALISM

Though modest global consequentialism is consistent with the value we commonsensically place upon the traits of character constitutive of

being a good friend, this turns out to be a pyrrhic victory. The reason is that modest global consequentialism cannot be distinguished in any principled way from a form of non-consequentialism. The real question, to which we will return in the next section, is thus whether *immodest* global consequentialism is consistent with the Friendly Justification Thesis, and the answer, to anticipate, is that it is not.

As we have seen, according to the modest global consequentialist the value of giving a slightly smaller benefit to a friend over a slightly larger benefit to a complete stranger is a *non-moral* value. There is, however, an alternative non-consequentialist understanding of the value of giving benefits to our friends. According to this non-consequentialist understanding, notwithstanding the fact that it is a relative value, the value associated with giving benefits to our friends is a form of *moral* value. The question we must therefore ask is whether there is any good reason to prefer the modest global consequentialist's view that the relative value associated with the giving of benefits to our friends is a form of non-moral value, to a form of non-consequentialism which is exactly like modest global consequentialism, except that it holds that the relative value associated with the giving of benefits to our friends is a form of moral value. In other words, is there any good reason to prefer a *non-moralized* interpretation of the Friendly Justification Thesis, which is what the modest global consequentialist insists upon, to a *moralized* interpretation, which is the interpretation favoured by the non-consequentialist?

Before answering this question it is important that we notice just how much the modest global consequentialist and his non-consequentialist opponent agree about. Both agree that, in a context in which an agent has to choose between giving a slightly smaller benefit to her friend or a slightly greater benefit to a complete stranger, but in which the consequences of her actions are otherwise identical, the evaluative significance, *simpliciter*, of her giving the smaller benefit to her friend is greater than that of giving the larger benefit to the complete stranger. It is not that (say) one thinks that there is *value* associated with the giving of the larger benefit to the complete stranger, whereas only *mere preference* is on the side of giving the smaller benefit to the friend. Nor is it the case that one thinks that while there is relative value at stake in giving the smaller benefit to the friend, such relative value as is at stake is outweighed by the neutral value associated with giving the larger benefit to the complete stranger, whereas the other thinks that the relative value at stake is of greater evaluative significance than the neutral value. They both agree that relative value is at stake and that it is of greater evaluative significance than the neutral value at stake. *All* that the modest global consequentialist and his non-consequentialist opponent disagree about is whether the value

associated with giving the smaller benefit to the friend is a moral or a non-moral value.

The issue, accordingly, is whether we can give any real content to this disagreement. In other words, assuming that there are both relative and neutral values, and assuming that the evaluative significance, *simpliciter*, of these values is much as the modest global consequentialist says they are, is there any determinate answer to the question whether the relative value associated with giving the smaller benefit to a friend is moral or non-moral in nature? In order to answer this question we must consider how the distinction between the moral and the non-moral is properly to be made.

#### A. Wolf on the Distinction Between the Moral and the Non-moral

Susan Wolf argues against the idea that moral standards are the ultimate standards by which to judge the value of the lives that people lead and the deeds that they do.<sup>12</sup> She claims that non-moral standards have a role to play in such judgements as well. Let's therefore see whether the distinction Wolf makes between the moral and the non-moral can help adjudicate the debate between the modest global consequentialist and his non-consequentialist opponent.

Wolf begins by considering three quite distinct views about the substance of morality that moral theorists standardly defend: utilitarianism, Kantian deontology, and common-sense morality. For each of these views about the substance of morality she then goes on to consider a range of examples of lives that people might lead, and deeds they might perform, which would be judged to be completely successful by the standards of the view in question. Having thus laid out these options she then proceeds to argue that people whose lives and deeds are judged to be completely successful by the standards of the three views can none the less have something good missing from their lives and that it would sometimes be better for them to have acted otherwise than as these views require.

In this way Wolf takes herself to show that there is a range of non-moral goods not captured by any of the views about the substance of morality that moral theorists standardly defend. The non-moral goods that she thinks aren't captured by the various moral theorists' views include being jovial, being athletic, being passionate about Victorian novels, being a dedicated oboe player, having the kind of sense of humour that would allow one to appreciate a Marx Brothers movie, being someone who enjoys preparing and consuming gourmet food like *paté de canard en croûte*, having Paul Newman's cool, and being

<sup>12</sup> Wolf, 'Moral Saints', *Journal of Philosophy*, lxxix (1982).

high-spirited and passionate like Natasha Rostov.<sup>13</sup> In Wolf's view the reason that standard moral views do not capture these goods is because they are not moral goods at all. They are rather personal goods, goods whose instantiation in our lives and deeds contribute not to 'moral perfection', but rather to 'individual perfection'.<sup>14</sup>

As is perhaps already plain, however, the distinction that Wolf makes between moral and non-moral goods is of no use at all in adjudicating the debate between the modest global consequentialist and his non-consequentialist opponent. For the goods Wolf brings to our attention are, to repeat, non-moral goods *no matter what our substantive moral view*. The relative value associated with giving a smaller benefit to a friend, rather than a larger benefit to a complete stranger, by contrast, is only non-moral by the lights of *certain* views about the substance of morality, views like that held by the modest global consequentialist. It is not non-moral by the lights of his non-consequentialist opponent's view about the substance of morality. By those lights the relative value is a moral value.

The upshot is thus that we need a much finer set of discriminations than Wolf provides us with. We need to know how, as among the various *competing* views about the substance of morality on offer – that is, as among theories none of which count as moral values being jovial, being athletic, being passionate about Victorian novels, and the like, but some of which count as moral the relative value of giving a smaller benefit to a friend, rather than a larger benefit to a complete stranger, and some of which don't – we are to decide which is the *correct* view about the substance of morality.

### *B. Railton on the Distinction Between the Moral and the Non-moral*

Peter Railton makes a somewhat different suggestion about the difference between the moral and the non-moral.<sup>15</sup> His account of the difference initially appears more helpful in adjudicating the debate between the modest global consequentialist and his non-consequentialist opponent.

According to Railton, the special character of moral evaluation, as opposed to non-moral evaluation, is properly captured by supposing that a moral point of view, as opposed to the non-moral point of view, is *impartial* in character.<sup>16</sup> Now this might well seem like a much more helpful suggestion as it might be claimed, as Railton himself in effect goes on to claim, that the neutrality of moral value in turn provides the

<sup>13</sup> *Ibid.*, 421 f.

<sup>14</sup> *Ibid.*, 436 f.

<sup>15</sup> Railton, 'Moral Realism', *Philosophical Review*, xcix (1986).

<sup>16</sup> *Ibid.*, 189.

best interpretation of the idea that the moral point of view is impartial.<sup>17</sup> If this were right then it would provide us with a decisive reason to prefer the modest global consequentialist's view to that of his non-consequentialist opponent. The relative value associated with giving a smaller benefit to a friend, rather than a larger benefit to a complete stranger, would then quite properly be regarded as a non-moral value because, being a relative value, it reflects a partial point of view, as opposed to an impartial point of view.

Tempting though the idea might be, however, it seems to me that we mustn't be too quick to agree with the suggestion that the neutrality of value provides the best interpretation of the claim that the moral point of view is impartial. After all, as Roderick Firth, who also thinks that the moral point of view is an impartial point of view, pointed out long ago, we can give the term 'impartial' a broader or a narrower interpretation.<sup>18</sup>

For example, according to Firth, Bentham's maxim that every man should count for one and none for more than one provides a natural interpretation of the idea of impartiality, but the interpretation is far too narrow if our interest in the idea of impartiality is to characterize what it is to have a moral point of view.<sup>19</sup> For if we were to accept this interpretation then it would be true by definition that a view like Ross's, according to which 'the rightness or wrongness of an act is determined in part by irreducible obligations arising directly from certain personal relationships', doesn't count as a moral view at all.<sup>20</sup> Yet, as Firth sees things, this would be completely unacceptable. For the truth or otherwise of a view like Ross's is 'synthetic and must not be prejudiced by our definitions'.<sup>21</sup> Ross's view thus is impartial, according to the broad interpretation of impartiality that Firth prefers. For similar reasons, it seems to me that we should therefore be alive to the possibility that a definition of impartiality in terms of the neutrality of value is also too narrow. Interpreted more broadly, the relative value associated with giving a smaller benefit to a friend, rather than a larger benefit to a complete stranger, would be regarded as impartial.<sup>22</sup>

As I understand him, however, Railton is quite aware of these

<sup>17</sup> *Ibid.*, 189–200.

<sup>18</sup> Firth, 'Ethical Absolutism and the Ideal Observer', *Philosophy and Phenomenological Research*, xii (1952).

<sup>19</sup> *Ibid.*, 337.

<sup>20</sup> *Ibid.*

<sup>21</sup> *Ibid.*

<sup>22</sup> What is the wider conception of impartiality that Firth has in mind? I take it that the wider conception of impartiality is something close to what Hare has in mind when he talks of universalizability. For more on universalizability, see the discussion of Hare in the next section.

dangers. For he describes himself as a providing 'reforming' definitions of moral rightness.<sup>23</sup> As he puts it, 'although the definitions proposed may not fit with all of our linguistic and moral intuitions, they nonetheless express recognizable notions of goodness and rightness'.<sup>24</sup> In other words, he admits that there is a certain indeterminacy in the bounds of the moral, and his suggestion that the moral point of view is an impartial point of view, where impartiality is in turn interpreted in terms of neutrality, is in effect a *stipulation*. But if this is right, then, far from Railton's account of the difference between the moral and the non-moral telling decisively in favour of modest global consequentialism, it in fact supports the idea that the concept of the moral is indeterminate between the modest global consequentialist's view and that of his non-consequentialist opponent. In the face of that indeterminacy the modest global consequentialist prefers a precisification of the idea of the moral according to which only neutral values count as impartial. His non-consequentialist opponent prefers a somewhat different precisification of the idea of the moral, one according to which relative values count as impartial. Their views are on all fours.

### C. Hare on the Distinction Between the Moral and the Non-moral

Perhaps the best-known attempt to formulate, in abstract terms, a clear and concise distinction between the moral and the non-moral is that made by R. M. Hare. Let's see, finally, whether his account of the difference tells in favour of the modest global consequentialist's view or that of his non-consequentialist opponent.

Hare's official focus is on judgements rather than values. The question that concerns him is, 'What makes a judgement a moral judgement, as opposed to a non-moral judgement?' In Hare's terms, then, the issue to be addressed is whether the judgement that it is desirable to provide a slightly smaller benefit to one's friends, rather than a slightly larger benefit to a complete stranger, in circumstances in which the consequences of one's actions are otherwise identical, is a moral judgement or a non-moral judgement.

Hare's suggestion is that the distinctive mark of moral judgements, as opposed to non-moral judgements, is that they are *prescriptive, universalizable, and overriding*.<sup>25</sup> He does not, of course, deny that there are non-moral judgements that are prescriptive, or that there are non-moral judgements that are universalizable, or that there are non-moral judgements that are overriding. He simply denies that any non-moral judgements satisfy all three conditions: satisfying all three

<sup>23</sup> Railton, 'Moral Realism', 204.

<sup>24</sup> *Ibid.*, 205.

<sup>25</sup> Hare, *Moral Thinking*, Oxford, 1981, p. 55.

conditions is, according to Hare, the mark of moral, as opposed to non-moral, judgements. Let's therefore ask, more specifically, whether judgements about the relative value of providing a benefit to a friend turn out to be moral, or non-moral, by Hare's standards.

Consider first the prescriptivity condition. To say that a judgement is prescriptive is, as I understand it, to say that anyone who accepts the judgement is motivated accordingly to some degree, at least absent practical irrationality.<sup>26</sup> But in that case, simply on the assumption that desirability judgements *as such* are prescriptive, it turns out that the judgement that it is desirable to provide a slightly smaller benefit to one's friends rather than a slightly larger benefit to a complete stranger will be prescriptive as well. There is, after all, nothing about this particular desirability judgement that would prevent it from being prescriptive when other desirability judgements are.<sup>27</sup>

Consider next the universalizability condition. To say that a judgement is universalizable is to say that someone who makes the judge-

<sup>26</sup> Smith, *The Moral Problem*, pp. 6 f.

<sup>27</sup> It is worth pointing out that we can provide a direct argument for the prescriptivity of judgements about the desirability of giving a slightly smaller benefit to a friend, over a slightly larger benefit to a complete stranger. I have argued elsewhere that judging it desirable that one acts in a certain way is, in effect, a matter of judging that one would want that oneself acts in that way if one had a set of desires that was maximally informed and coherent and unified (Smith, *The Moral Problem* and 'In Defense of *The Moral Problem*: A Reply to Brink, Copp and Sayre-McCord', *Ethics*, cviii (1997)). If this is right, however, then it follows that those who make the judgement that it is desirable to provide a slightly smaller benefit to their friends rather than a slightly larger benefit to a complete stranger in effect judge that they would want that they themselves provide a slightly smaller benefit to their friends rather than a slightly larger benefit to a complete stranger if they had a set of desires that was maximally informed and coherent and unified. But those who make this judgement would then be required, by norms of coherence, to desire to provide a slightly smaller benefit to their friends rather than a slightly larger benefit to a complete stranger. They would be required to have this desire because the belief-desire pair that comprises the belief that one would want that one gives a slightly smaller benefit to one's friend rather than a slightly larger benefit to a complete stranger, on the one hand, and the preference to give one's friend that smaller benefit on the other, is a more coherent pairing of belief and desire than that which comprises the belief together with indifference or aversion to providing one's friend with that smaller benefit. And, if this is right, then it follows immediately that, absent practical irrationality – that is, absent incoherence in their psychology – those who make the judgement about the desirability of giving a smaller benefit to their friends would indeed desire accordingly. The judgement is thus prescriptive. (Note that this amounts to a more precise suggestion about the nature of evaluative significance, *simpliciter* (Michael Smith, 'Normative Reasons and Full Rationality: Reply to Swanton', *Analysis*, lvi (1996), and Smith, 'In Defense'). Neutral value is a matter of what would be desired if we had a set of desires that was maximally informed and coherent and unified where the desires in question have non-indexical content. Relative value is a matter of what would be desired if we had a set of desires that was maximally informed and coherent and unified where the desires in question have indexical content. Evaluative significance, *simpliciter*, is a matter of the relative strength of the non-indexical and indexical desires that we would have if we had a set of desires that was maximally informed and coherent and unified.)

ment is logically committed to making the same judgement about any case that is identical in terms of its universal features. But someone who judges it desirable to give a slightly smaller benefit to his friends rather than a slightly larger benefit to a complete stranger, as opposed to someone who merely has an arbitrary preference to give the benefit to that particular individual (who just so happens to be his friend), is indeed logically committed to judging it likewise desirable to prefer giving a slightly smaller benefit to another of his friends, rather than a larger benefit to a complete stranger, in all of those possible cases in which the particular individuals involved are replaced with others, but in which the relations between the individuals are maintained. For in all such cases the person must decide on the relative merits of providing a slightly smaller benefit to a *friend* – a different friend, of course, but still a friend and a friend of exactly the same standing – as opposed to a slightly larger benefit to a complete stranger, albeit an exactly similar stranger. The judgement would therefore seem to be universalizable.

Finally, consider the overridingness condition. As I understand it, to say that a judgement is overriding is to say that someone who makes that judgement judges the value so judged to be of greater significance, *simpliciter*, than any other values of which they judge. Needless to say, it thus turns out that, at least by the lights of both the modest global consequentialist and his non-consequentialist opponent, the judgement that it is desirable to give a slightly smaller benefit to a friend, rather than a larger benefit to a complete stranger, is overriding as well. For they both agree that this is what the agent has most reason to do.

The upshot is thus that, at least by the lights of the modest global consequentialist and his non-consequentialist opponent, the judgement that it is desirable to provide a slightly smaller benefit to a friend, rather than a slightly larger benefit to a complete stranger, satisfies all three of the conditions Hare specifies. If we accept Hare's account of the distinction between the moral and the non-moral, then, as between the two theories on offer, it is the non-consequentialist's claim that the judgement that it is desirable to give a slightly smaller benefit to a friend, rather than a larger benefit to a complete stranger, is a moral judgement, as opposed to the modest global consequentialist's claim that it is a non-moral judgement, that is the more plausible. Needless to say, this result will be of no comfort to the modest global consequentialist.

But, I think, the modest global consequentialist should not be too worried by it. Hare's claim that moral judgements are overriding overstates the significance of morality. This is not, of course, to say that moral values can be completely insignificant. On the contrary, as

Robert Adams notes, 'it is commonly', and I would say rightly, 'made a criterion for a theory's being a theory of *moral* obligation, that it claim a special seriousness for its judgements of obligation'.<sup>28</sup> But this falls far short of the idea that moral judgements are overriding. It requires only that moral values have *significant* weight. They cannot be dismissed as *trivial* considerations in a choice situation. Even though they can be overridden, then, they cannot be overridden lightly.

Indeed, the fact that Hare has overstated the case is made even more apparent when we remember Wolf's discussion of the difference between the moral and the non-moral. As you will recall, Wolf specifies a whole range of values such as being jovial, being athletic, being passionate about Victorian novels, and the like, that do not count as moral on any of the competing views about the substance of morality that people standardly defend. I take it that Wolf thinks that these must therefore turn out to be non-moral on any ultimate conception of morality. But this provides us with sufficient reason to suppose that Hare is wrong to claim that moral judgements are overriding. For Wolf's whole point is that some of the non-moral judgements that she focuses on sometimes override moral judgements. Since these non-moral judgements will almost certainly satisfy Hare's other two conditions – that is, since they will be both universalizable and prescriptive – it follows that we therefore need to weaken the overridingness condition.

#### V. A SECOND ATTEMPT AT STATING THE CHARGE AGAINST CONSEQUENTIALISM

It seems to me that the only conclusion to draw from the preceding discussion is that the relative value associated with giving benefits to a friend rather than a complete stranger does not fall clearly within the bounds of the moral or the non-moral. There is a certain indeterminacy in our concept of the moral, and the relative value associated with giving benefits to a friend rather than a complete stranger falls right within that area of indeterminacy.

If this is right, however, then it follows that there is no serious debate to be entered into between a modest global consequentialist, on the one hand, and a non-consequentialist, on the other, whose only disagreement with the modest global consequentialist is about whether the relative value associated with giving benefits to a friend rather than a complete stranger is moral. Both parties to this debate erect their moral views on the basis of an untenable stipulation about where the boundary between the moral and the non-moral lies. Consequen-

<sup>28</sup> Adams, 'Motive Utilitarianism', *Journal of Philosophy*, lxxiii (1976), pp. 477 f.

tialists thus cannot take refuge in a modest global consequentialist account of the value of the traits of character constitutive of being a good friend and the acts that good friends are disposed to perform. We must therefore ask what the global consequentialist should say about the value of these traits of character and acts.

The global consequentialist needs some principled reason for supposing that the relative value associated with giving benefits to a friend rather than a complete stranger is non-moral in nature. The discussion of Wolf, Railton, and, most importantly, Hare in the last section suggests that there is only one way in which he can do this. He needs to rely on a supplementary claim about the weight of relative value as against neutral value. Specifically, he must suppose that relative values do not have significant weight, as compared with neutral values.

Let's call this view, for obvious reasons, *immodest global consequentialism*. The immodest consequentialist holds not just that neutral values are the sole determinant of moral rightness, but also that neutral values are, as such, of much greater evaluative significance, *simpliciter*, than relative values. It would seem that there are various reasons an immodest global consequentialist might have for thinking that this is so. One possibility is that the immodest global consequentialist thinks that there are no relative values to be weighed on the scales against neutral values to see which has greater evaluative significance, *simpliciter*.<sup>29</sup> Another possibility is that, though he thinks that there are relative values to be weighed on the scales, he thinks that no amount of relative value, no matter how great, could ever have more evaluative significance, *simpliciter*, than even the smallest amount of neutral value. In other words, he thinks that neutral values are lexically prior. Another possibility is that, though he does not think that neutral values are lexically prior, the only contexts in which he admits that relative values outweigh neutral values are contexts in which the relative value is massive and the neutral value is minute.

Whatever the immodest consequentialist's particular reasons for denying that relative values have significant weight, the upshot is that he will most certainly deny that, in situations in which the consequences of our actions are otherwise identical, the relative value associated with giving a slightly smaller benefit to our friends has greater evaluative significance, *simpliciter*, than the decrease in neutral value associated with failing to give a slightly greater benefit to a complete stranger. It will therefore come as no surprise to hear that the immodest global consequentialist will have a hard time squaring his theory with our most deeply held convictions about the

<sup>29</sup> Jackson.

value of the traits of character constitutive of being a good friend and the acts that good friends are disposed to perform.

Consider once again the Friendly Justification Thesis. It tells us that among the traits of character possessed by a good friend we find the disposition to believe that, in choosing between alternatives that are otherwise identical, the increase in relative value associated with the provision of a smaller benefit to one of her friends is of greater evaluative significance, *simpliciter*, than the decrease in neutral value associated with failing to provide a greater benefit to a complete stranger. According to the immodest global consequentialist, however, though the possession of this belief maximizes neutral value, the belief itself is straightforwardly false. To repeat, the mismatch with our most deeply held convictions couldn't be more apparent. For we are convinced that the belief is true.

## CONCLUSION

I have argued that the fact that we place the value we do on the traits of character constitutive of being a good friend, and the acts that good friends are disposed to perform, creates a considerable problem for immodest global consequentialism. This is bad news because, as I have also argued, immodest global consequentialism is the only version of consequentialism that can be distinguished in any principled way from a form of non-consequentialism.

The charge against the immodest global consequentialists is, in essence, that the very best they can do by way of vindicating our most deeply held convictions about the value of these traits of character and actions isn't good enough. For the immodest global consequentialist's attempted vindication of the value of the traits of character that good friends possess undermines the truth of convictions we have about the value of the acts that good friends perform. How serious this charge against immodest global consequentialism is, however, is a matter of some legitimate dispute.

For example, one consequence is that, absent self-deception or gross irrationality, no one who has the beliefs that immodest global consequentialism tells him he morally should have could simultaneously believe the truth of immodest global consequentialism, as this would in effect be to have contradictory beliefs. It would require that he believe both that relative values are of greater evaluative significance than neutral values and that neutral values are of greater evaluative significance than relative values. I take it that the reason Derek Parfit argues that the mere fact that a theory is self-effacing – that is, the mere fact that a theory recommends its own rejection (which is what immodest global consequentialism does) – is no objection to that

theory is *precisely* because he wants to play down the significance of this consequence.<sup>30</sup>

But the important question, I think, is not whether the very idea of a theory's being self-effacing is incoherent. Rather, granted that that idea is not incoherent, the important question is whether we could have any good reason to believe a self-effacing moral theory when the reason that it is self-effacing is that, on the one hand, it morally requires us to have some of the deeply held normative convictions that we do, and, on the other, it is itself blatantly inconsistent with the truth of those normative convictions. I must confess that I can think of no such reason. But before jumping to the conclusion that consequentialism is therefore false, perhaps we should encourage the consequentialists to speak for themselves.<sup>31</sup>

msmith@coombs.anu.edu.au

<sup>30</sup> D. Parfit, *Reasons and Persons*, Oxford, 1984, p. 24.

<sup>31</sup> I would like to thank Frank Jackson for helpful comments on an earlier version of this paper.