CANADIAN JOURNAL OF PHILOSOPHY
Volume 17, Number 3, September 1987, pp. 643-654

*Donagan On The Sins Of Consequentialism**

SHELLY KAGAN
The University of Illinois at Chicago
Chicago, IL 60680
U.S.A.

Most intuitively forceful criticisms of utilitarianism, I believe, reduce to two basic objections. Both arise from the relentlessness of the utilitarian injunction to promote the overall good. On the one hand, this means that agents are permitted to perform an act of any kind whatsoever—provided only that the consequences of that act are better than those of any alternative. In particular, this means that it is permissible to impose tremendous sacrifices or injuries upon someone, if this is the only way to ensure even greater gains for others. But in allowing agents to deliberately impose harm on someone for the sake of others, utilitarianism seems to permit acts which are morally unacceptable. Intuitively, we want to say that rights, or other moral prohibitions, forbid treating people in certain ways—ways allowed by utilitarianism. This, then, is the first major objection: utilitarianism permits too much.

On the other hand, the utilitarian injunction demands that the agent promote the overall good to the limit of his abilities. Agents must push themselves to make their maximal contribution, no matter what the cost to themselves might be—provided only that their own sacrifice is outweighed by gains to others. In effect, there is no limit to the demands which utilitarianism places upon the individual; and the required sacrifices are not merely those of money, effort, or risk of bodily injury. For utilitarianism leaves the agent no room to favor those things he holds dear: e.g., it forbids him to give more weight to the welfare of those he loves than they would merit from an impersonal

* This paper was presented at a conference of the Illinois Philosophical Association on 'The Philosophy of Alan Donagan,' November 4, 1983.

standpoint; and it requires that he abandon his personal projects whenever, and simply because, doing so would have better consequences. Intuitively, however, we believe that there *is* a limit to what morality can demand of us; we are not morally required to be forever promoting the overall good – not even within permissible means. Some acts, although meritorious, are above and beyond the call of duty. This, then, is the second major objection: utilitarianism *demands* too much.

There are, of course, other common objections to utilitarianism, and several of these are discussed by Alan Donagan in *The Theory of Morality*.[1] For example, there are problems of coordination and cooperation (195-6); issues relating to publicity, and whether utilitarianism supports its own rejection (197-9); and the question of our limited ability to predict the consequences of our acts (199-205). These are important theoretical issues for a utilitarian to come to grips with; but I do not think that they genuinely *trouble* anyone. Ultimately, clever technicians either will, or will not, be able to produce adequate solutions to these problems. But I suspect that *no one* has ever become a utilitarian, or rejected utilitarianism, because of their own views about these questions.

The matter is considerably different with the two basic objections I have mentioned. These seem quite forceful indeed. Utilitarians are generally embarrassed by these ethical implications of their theory; and deontologists seem to get closer to the heart of what is *wrong* with utilitarianism when they argue that it permits too much and demands too much. These objections are central, because it *matters* that utilitarianism gives the wrong answers: people turn to an ethical theory for guidance, and it is morally objectionable if the theory misguides them. There may be other 'problems' with utilitarianism, but permitting the impermissible and demanding the nonobligatory – these are the *sins* of utilitarianism.

Donagan himself seems to have been aware of this when he wrote his earlier paper, 'Is There a Credible Form of Utilitarianism?'[2] It is true that in this essay Donagan was criticizing a particular version of rule-utilitarianism, while we are here considering unabashed act-utilitarianism. But since the general thrust of his article seems to be that even rule-utilitarianism cannot totally escape crippling objections of the sort that plague act-utilitarianism, I hope it is not unfair to ascribe

---

1 Alan Donagan, *The Theory of Morality* (Chicago: University of Chicago Press 1977). All page references in the text are to this work.

2 Alan Donagan, 'Is There a Credible Form of Utilitarianism?' in Michael D. Bayles, ed., *Contemporary Utilitarianism* (New York: Doubleday 1968)

to him the view that his important objections to the former are, a fortiori, important objections to the latter.

In that article, Donagan offered two intuitive considerations against the version of rule-utilitarianism he examined—and they correspond to the two basic objections I have described. His first criticism was that rule-utilitarianism misclassifies certain supererogatory acts as being obligatory. This, of course, is an instance of our more general objection that utilitarianism demands too much—requiring sacrifices which morality does not actually require. His second criticism was that rule-utilitarianism inevitably permits at least some cases of harming innocent men for the sake of the public good. And this, obviously, is an example of our general objection that utilitarianism permits too much—here, violating fundamental rights of the innocent. I am suggesting, in effect, that it is not mere, happy coincidence that Donagan offered these two objections: rather, these are the points at which the deontologist's case against utilitarianism either stands or falls.

It is worth noting, however, exactly which feature of utilitarianism is attacked by our two objections. Utilitarianism is the combination of two views: a theory about what makes acts right or wrong; and a theory about what makes outcomes better or worse. Utilitarianism's theory of the right is *consequentialism*, which says, roughly, that an act is permissible if and only if it has better consequences than any alternative act. This view, obviously enough, is compatible with any view whatsoever about what factors make one outcome better than another. The particular theory of the good endorsed by utilitarians says that *individual well-being* is the only factor which affects the value of an outcome. This view, in turn, is compatible with alternative accounts of what makes an act right.

Now it might well be objected that this theory of the good neglects many other factors which can affect the value of an outcome;[3] and many who call themselves 'utilitarians' have some sympathy with this complaint. But we can put this aside—for it is its maximizing theory of the right—consequentialism—which accounts for the fact that utilitarianism permits and demands too much. On *any* plausible theory of the good, there will be cases where the only way to promote the good would be to harm some innocent individual—and consequentialism will always permit such acts. And there will inevitably be cases where promoting the good would require tremendous sacrifices from the agent—and consequentialism will demand such constant pursuit of the good. Essentially, then, utilitarianism gives the wrong answers because of

---

3 See, e.g., the example mentioned in note 6 below.

its consequentialist basis. The sins of utilitarianism are really the sins of consequentialism.

I have said that our two objections are intuitively forceful: they appeal to vivid intuitions about what one may or may not do—and they sharply bring out how consequentialism runs afoul of these intuitions. Were this sufficient to establish the validity of the objections, deontology's victory would be assured. But such appeal to intuition is *not* sufficient. Intuitions maybe the first word in moral theory; but they are certainly not the last—for intuitions may be corrupt or simply mistaken, and even deeply entrenched intuitions are not sacrosanct. Thus an adequate philosophical defense of a moral theory must go beyond the mere appeal to intuitions. Indeed, it may even lead us to *reject* some of our initial intuitions.

Donagan certainly recognizes the need to offer a grounded defense of the particular deontological theory he supports. *The Theory of Morality* is a sustained attempt to derive the details of what he calls 'common morality' from a Kantian 'fundamental principle,' and then in turn to offer support for the Kantian principle itself. In staking out a position so clearly opposed to consequentialism, we would expect Donagan to take great pains to establish that the *true* morality—unlike consequentialism—forbids certain acts, even though they might promote the good, and does not demand of agents the relentless pursuit of the good, not even with permissible means. I want to suggest, however, that Donagan fails to establish either of these claims.

Let's start with the objection that consequentialism permits too much. Making the plausible assumption that human well-being is at least *one* relevant factor in evaluating outcomes, imagine that unless I kill one innocent person, two other innocent persons will be murdered—and there are no other relevant consequences. (In real life, of course, situations are never this clear and simple. But unless there really could be cases where, e.g., harming someone has better consequences, then deontology and consequentialism cannot here disagree.) Killing the one to save the two has better consequences than refraining. Does Donagan prove that killing the one is nonetheless morally forbidden?

Donagan takes the fundamental principle of morality to be: 'It is impermissible not to respect every human being, oneself or any other, as a rational creature' (66). Granting him this, can we derive a prohibition against killing the one? Donagan argues that

> Since respect for human beings as rational creatures entails, in general, treating every normal adult as responsible for the conduct of his own affairs, to interfere by force with anybody else's conduct of his life, unless there is a special and adequate reason, is not to respect him as a rational creature. The principle may therefore be laid down that *it is impermissible for anybody at will to use force upon another*.(82)

From this general prohibition, Donagan quickly derives the precept that 'no man may at will kill another' (83). This all seems plausible enough (although I think it might be challenged). Does it show that I may not kill the one to save the two? No, it does not show this—for the whole question, of course, is, what if any are the permissible exceptions to the general prohibition against killing? Donagan, in the passage just quoted, recognizes that there may be 'special and adequate reasons' for killing someone—in which case killing is not failing to show respect. The consequentialist can plausibly maintain that he does not kill 'at will'—but only in those special cases where doing so, e.g., avoids even greater suffering. Is this a legitimate exception to the generalization?

It is Donagan's own approach to argue for a general principle, and then mark off possible exceptions (72-4). True to form, after offering the prohibition against killing at will, he notes that 'there are circumstances in which it is legitimate to kill or injure others' (83). Self-defense and defending others are two such legitimate exceptions (84-5), and these in turn suggest the possibility of just wars (87); capital punishment is also mentioned, in passing (87; also 163). But are there other exceptions? This question is never raised. Yet what Donagan *should* have done is give some reason to think that he has stated the only major exceptions. In particular, Donagan needs to argue against the consequentialist that killing the one to save the two—or, more generally, harming one for the sake of others—is *not* a legitimate exception. But this is a problem never directly faced; it is buried in the discussion of a different point.

After arguing that defending others is permitted, Donagan goes on to insist that it is actually *required*—subject to two qualifications. This follows from the principle of beneficence, which states: 'It is impermissible not to promote the well-being of others by actions in themselves permissible, inasmuch as one can do so without proportionate inconvenience' (85). The claim that one needn't promote the well-being of others when this involves disproportionate inconvenience obviously relates to the issue of how demanding morality is—and we shall come back to this. What about the requirement that the well-being of others be promoted only by actions 'in themselves permissible'? What does this condition amount to?

Donagan writes that

> it is absolutely impermissible to promote the well-being of others by any action that is impermissible in itself. This condition is implicit in common morality as a system (cf. 5.2). Accordingly, it is impermissible to secure a benefit for some (say, a legacy that will enable a physician to open a free clinic for the poor) by doing wrong to others (say, by killing the testator who has bequeathed that legacy). (86)

This is a fascinatingly obscure passage. The example of killing the wealthy benefactor makes it clear (if we harbored any doubts) that Donagan does think harming the innocent for the sake of others is forbidden. What is not at all clear is how this is supposed to follow from anything he has told us—even in this passage. It is a substantive conclusion indeed that one may not promote the well-being of some by harming others. But this is supported only by trivial assertions. After all, to observe that the well-being of others may not be promoted by an act that is 'impermissible in itself'—is apparently to say nothing more than that we must not promote the good through *forbidden means*. But even the consequentialist accepts this platitude: we must never perform forbidden acts. The controversial issue, however, is whether harming one for the sake of others *is* such a forbidden act—and the platitude cannot help us here.

Similarly, we can grant Donagan that it follows that we may not secure benefits for some 'by doing wrong' to others—for this simply restates the platitude. But is killing the benefactor doing *wrong* to him? It is, if it is forbidden to kill him—and Donagan clearly *believes* that it is wrong: but he has given us no reason to think so. It is easy to slide from the trivial claim that we must not harm someone wrongfully, to the substantive claim that harming him is wrong. But the slide is an unjustified one nonetheless. It appears that Donagan simply begs the question against the consequentialist, and makes an ad hoc assumption about what follows within his own system.

Perhaps I have misread the passage. Donagan refers us to a later section in the book (5.2), and perhaps this earlier passage simply assumes the substantive claim—which will not be defended until that later section. Unfortunately, I find the later passage no more helpful. After discussing the meaning of the Kantian distinction between perfect duties (like those based on the prohibitory precept against killing at will) and imperfect duties (like those derived from the principle of beneficence), Donagan claims that 'it should now be plain' why precepts requiring the promotion of human well-being

> contain the condition, "by morally permissible actions" (cf. 3.2, 3.3). A fundamental principle which categorically forbids violating the respect owed to human beings as rational must condemn any plan for promoting human well-being by which that respect would be violated. Hence the fundamental principle of morality itself entails that every precept ... ordaining the promotion of human well-being generally—shall contain as a condition that ... no prohibitory precept—may be violated.
>
> And that is the true sense of the Pauline principle that *evil is not to be done that good may come of it*. The evil that is not to be done is the violation of the prohibitory precepts ...; and the good for which it is not to be done, even though common morality requires that it be promoted, is the well-being of oneself and others as human. (154-5)

This passage strongly suggests that Donagan thinks he has established a substantive conclusion—i.e. that we may not harm in order to bring about good consequences.[4] But all the passage actually tells us is the same old platitude. That we may not *violate* a prohibitory precept for a good end, is uncontroversial enough; after all, we may not violate *any* valid precept, period. The question at issue, however, is whether or not killing the one to save the two, e.g., *is* a violation of the prohibitory precept against killing at will—or whether it is a legitimate exception, as the consequentialist would have it. It is, once more, easy to slide from the trivial observation that evil—i.e. *wrongdoing*—is not to be done, to the substantive claim that *harm* is not to be done. But it is a non sequitur for all that—and Donagan offers us nothing better.[5]

Very well then, Donagan fails to prove that consequentialism permits too much. But isn't it a simple matter for us to provide the proof ourselves? Harming someone for the benefit of others fails to respect him; and since the fundamental principle requires respect, such acts are forbidden.

Such an argument, however, is inadequate. It is easy to *assert* that killing the one to save the two fails to respect the one—but I do not see how to *defend* this assertion. It will not do to say that it is self-evident, for surely the sincerity of consequentialists should convince us that it is not evident to them. Indeed, the consequentialist can say more. He can insist that I am *not* failing to respect the one—even though I kill him: for the well-being of the one was counted just as much as that of the others. The deontologist thinks this view ludicrous, perhaps; but we agreed long ago that appeal to intuition is insufficient. And I simply do not see how the deontologist can *support* his view about what respect entails.

Let me turn now to the second objection—that consequentialism *demands* too much. Here I can be somewhat more brief. In defense of the principle of beneficence, Donagan claims that 'if a man respects other men as rational creatures' he will 'further their efforts as far as he prudently can' (85). Although the egoist might doubt whether

---

4 It may be that Donagan's claim that 'it should now be plain' why the precepts contain the condition 'by morally permissible actions' indicates a belief that the matter has, in effect, just been explained. But I can find nothing in the immediately preceding paragraphs to support such a view. It is also possible that the quoted passage takes the substantive claim to have been already established in the earlier sections to which it refers; but it would be circular to have each passage appealing to the other.

5 Donagan also offers a schematic version of his argument (157); but this appears to embody the same mistake.

respect ever entails a requirement to aid at all—surely the consequentialist will not challenge this claim. But what is the extent of the obligation to aid? Recall the second of Donagan's two qualifications to the principle of beneficence: one must promote the well-being of others, but only 'inasmuch as one can do so without proportionate inconvenience.' What does this qualification come to? Donagan's sole explanation is that

> One does not fail to respect another as a rational creature by declining to procure a good for him, if that good can be procured only by relinquishing an equal or greater good for oneself. Hence, ... to promote the well-being of others at the cost of one's life or fundamental well-being would be supererogatory. (86)

On what I take to be the most plausible reading of this passage, Donagan is saying that a sacrifice is disproportionately inconvenient exactly when it is equal to or greater than the potential gain to others. If so, it seems to me that this passage implies that Donagan's system is quite as demanding as consequentialism. This is rather surprising, for Donagan certainly gives the *impression* that he has defended the traditional view that the principle of beneficence is fairly modest, all told: not requiring constant sacrifice, and leaving the agent considerable freedom to pursue personal projects and to favor loved ones—even though such acts may well fail to make the agent's maximal possible contribution. Donagan speaks, if only in passing, of the permissibility of 'declining an unreasonable burden' (126), and of an act's being nonobligatory 'because it demands too much of the agent' (156). (Also cf. 78, 156, and 169.) Although he never explicitly says so, I find the suggestion irresistible that Donagan *thinks* his principle of beneficence is reasonably modest in its demands. If this is Donagan's view, however, it is simply mistaken.

For if we are required to promote the well-being of others except when this involves a disproportionate inconvenience to ourselves, and if a sacrifice is disproportionate only when it is equal to or greater than the potential gain to others, then we are required to promote the well-being of others—no matter what the cost to ourselves—provided only that the *overall* gains outweigh the overall losses. Donagan claims that promoting the well-being of others at the cost of my own life would be supererogatory. But this overlooks the fact that by sacrificing my life or fundamental well-being I may be able to save the lives of *several* others; and if so, it cannot be claimed that my sacrifice is disproportionately inconvenient. Indeed, the opportunity to save the lives of others (e.g. famine victims) through sacrifices of time and money is a constant one—and beneficence apparently demands that I make these sacrifices.

But this is exactly what consequentialism demands: we must promote the overall good as much as we possibly can, remembering only that our own good is to be counted too. It may be that consequentialism is too demanding—but Donagan has given us no reason to believe this claim.[6] What he has given us is reason to believe that common morality is just as demanding as consequentialism.

Perhaps, however, I have misunderstood what Donagan means by 'disproportionate inconvenience.' Strictly speaking, in the passage quoted all that Donagan explicitly commits himself to is the position that a sacrifice is disproportionately inconvenient when it is equal to or larger than the gain to the other; he does not explicitly say that it is disproportionate *only* in such cases. It is thus compatible with a literal interpretation of this passage to hold that a sacrifice can also be disproportionately inconvenient even in some cases where the cost is *less* (perhaps even much less) than the gain to others. If so, then the principle of beneficence will indeed be less demanding than consequentialism.

I find this alternative interpretation unlikely. After all, the passage I have quoted is the only place where Donagan explains what he means by disproportionate inconvenience; nowhere does he express the rival view just suggested. I find it puzzling why Donagan would not have taken the opportunity (in the quoted passage, or elsewhere) to tell us the *full* range of what is to count as disproportionate inconvenience. (On the original interpretation, in contrast, Donagan did just what one would expect—indicating exactly which cases involve disproportionately inconvenient sacrifices.) Furthermore, if a case can involve disproportionate inconvenience even though the sacrifice is less than the gain to others—how *much* less? The alternative interpretation simply has to insist that Donagan inexplicably fails to address this obvious and pressing question. Most importantly, if we adopt the alternative interpretation, we should be struck by the fact that Donagan offers no *defense* for the view now being ascribed to him—i.e., that in some cases a sacrifice need not be made even though it involves less cost to the agent than the potential gain to others.

---

6 Donagan does criticize utilitarianism for requiring that we make sacrifices to aid the lazy and the wicked who have brought their woe 'upon themselves' (209). To the extent that this objection is plausible, however, I think that it turns on the fact that we don't really feel that the consequences are better when sacrifices are made for the undeserving (even if there *is* a net gain in mere happiness). Such judgments about desert could be incorporated into a theory of the good. If a consequentialist theory of the right were combined with such a theory of the good, it would escape the counter-example.

Donagan may have failed to defend a modest principle of beneficence—but again we ask, can't we provide the defense ourselves? The fundamental principle requires respect; perhaps, however, respect only requires that we give the well-being of others *significant* weight—but not necessarily as much weight as we give our own well-being or that of those we love. If this is so, then we may permissibly refuse to take on certain sacrifices, even though the gains to others would be much greater.

Once again, however, the crucial premise is difficult to defend. The consequentialist may object that respecting others involves valuing them as highly as we do ourselves—counting their interests no less than our own. Many deontologists will insist against this that respect for all is compatible with taking a special interest in one's own well-being. This may even seem obvious. But inadequate appeal to intuition aside, I simply do not see how the deontologist might *defend* his view about what respect entails.

We seem to have reached an impasse: we are unable to settle how much morality should permit, and how much morality should demand. I want to conclude by quickly suggesting what brought us to this state. Donagan, along with a good many others, has attempted to ground morality in the notion of respect: actions are right, or wrong, *because* they respect, or fail to respect, persons.[7] But this seems to me to be exactly backwards. With rare exceptions, it is only *by virtue* of its being morally right or wrong that an action shows respect or disrespect. Ultimately, I think, to show respect for individuals is to regard them as valuable beings and to treat them in the manner appropriate for beings of this sort. But this means that we cannot decide whether a given act is respectful until we have *first* determined how it is appropriate to treat beings of this sort. And this is precisely what morality must tell us.

If I am right about this, we cannot hope to settle genuine disputes about morality by appealing to judgments about respect—for such judgments inevitably appeal (often covertly) to judgments about morality. Such arguments from respect, then, reduce to little more than beggings of the question—and both sides, of course, can play at that game.

---

7 In informal comments prepared in initial response to this paper, Donagan suggested that he did not attempt to ground morality on the notion of 'respect' or 'respect for persons' per se, but rather on the notion of 'respect for persons as rational creatures' (a concept explicated especially at 59-66). This is certainly true, but I do not believe it would weaken any of my arguments if we were, cumbersomely, to substitute the third expression where I have used one of the first two.

Thus, it is the very attempt to ground morality in respect which has brought us to the present impasse. If we are going to make genuine progress in settling the disputes between deontology and consequentialism, we must take a different route.

Received October, 1985
Revised April, 1986