

Deontology and Alan Donagan's problem of exception-rules

JOHN DAVENPORT

1. Introduction

In his *The Theory of Morality*,¹ Alan Donagan attempts to systematize and clarify the moral conception found in the 'common code' or 'moral law' of the Hebrew-Christian tradition (p. 6). As part of this project, Donagan attempts to defend the position of 'common morality' that the scope of its moral precepts should not be limited by purely 'consequentialist' considerations. Donagan's recognition of this problem has interesting general implications, because it reveals the need for deontological theories of morality to derive principles for determining the scope of precepts *from their own grounds*. I will show that Donagan tries to solve this problem by relying implicitly on a two-part standard to determine whether a limitation on the scope of a moral precept is justified. However, both parts of Donagan's standard turn out to be inadequate: the first because it is incompatible with Donagan's own deontological approach, and the second because its implications are more ambiguous than Donagan realizes.

The system Donagan sets out in *A Theory of Morality* is based on a Kantian fundamental principle categorically requiring respect for human beings as rational. The resulting theory contains (at the first-order level) 'prohibitory precepts and precepts commanding the promotion of certain ends' (p. 153), as well as a rule governing their relation. We might distinguish these as follows: (1) certain perfect duties to oneself and others requiring us 'not to do, or not to omit, a certain kind of action'; (2) certain imperfect duties, such as the principles of 'culture' and 'beneficence,' which require us 'to promote a certain general end' (p. 154); and finally (3) the 'Pauline principle,' which implies that the perfect duties are always *absolutely prior* to the imperfect ones:

the fundamental principle of morality itself entails that every precept of imperfect duty following from it – every precept ordaining the promotion of human well-being generally – shall contain a condition that no precept of perfect duty following from it – no prohibitory precept – may be violated. ... that is the true sense of the Pauline principle that *evil is not to be done that good may come of it*. (p. 155)

¹ Alan Donagan, *The Theory of Morality* (Chicago and London: University of Chicago Press, 1977). All page references to this text are given parenthetically within the article.

As a result of this Pauline Principle, Donagan maintains that certain basic moral prohibitions, whose observance is essential to the moral rightness of any action, are prior to moral principles requiring the promotion of certain general goods that contribute to well-being. This kind of priority between classes of duties, which I will call a *priority in the application of principles*, is often confused with the so-called deontological 'priority of the right over the good', which orders considerations used to *justify* moral principles. The latter deontological priority requires that (1) categorical ('perfect') requirements and prohibitions of certain kinds of actions, (2) 'imperfect' duties, (3) the scope of these moral precepts, and (4) the proper weighting relations between the different classes of precepts, are all to be determined from a fundamental principle (or principles) of right that involve something other than unmediated intuitions and mere summations of preference-satisfaction.² This first principle is usually a formalist one, although it is not evident that it necessarily must be.

This distinction between priority in application of principles and deontological priority in justification suggests that consequences may be relevant for moral determinations in deontological theories, as long as their relevance is determined by a principle of moral legitimation which specifies the limit of the morally acceptable in terms of something other than *consequences alone*. For example, according to Donagan, the 'principle of the least evil', which takes account of consequences, comes into play in judgments regarding 'the formation of rational plans under the precepts of imperfect duty, when there is a choice between unequal goods' (p. 155) and also possibly in determining the right course of action where questions of double-effect are concerned (p. 161). Consequences of actions *always* have serious moral significance on Donagan's account, since he holds that 'an action is identical with the causing of each and every consequence to which the doer's agency extends' (p. 160).

It is crucial, however, to distinguish the open-ended allowance that consequences may play a role in moral deliberations from *consequentialism* in the narrower sense of the term, as used by Donagan and Anscombe. **With them, I will reserve 'consequentialism' in the strict sense for the view** that rules are justified as 'binding absolutely and unconditionally' solely on the basis of the consequences of observing or violating those rules

² For example, in John Rawls's theory of justice, the basic deontological ordering of considerations determines that substantive principles of justice are to be justified by selection (in a stage-wise fashion) within the Original Position including its Veil of Ignorance. The principle determining the order of application among substantive principles, however, is the lexical priority of the formal liberties over the principles guaranteeing social and economic justice. This specific principle of priority in application is thus not essential to deontology *as such*.

(p. 173).³ Strict consequentialism thus limits the possible range of relevant considerations in a way that deontological theories need not.

In *The Theory of Morality*, Donagan is resolute in attacking theories that are 'consequentialist' in this strict sense, and in defending the scope of traditional precepts in 'common morality' against 'consequentialist' objections. In the next section, I will describe the two-fold standard Donagan employs in this defence of traditional moral precepts. In §3, however, I will argue that the first of Donagan's two criteria itself turns out to be 'consequentialist' in the objectionable sense. Finally, in §4, I will argue that Donagan has wrongly assumed that his second, 'epistemic' criterion can only be used to defeat arguments for exceptions to moral precepts. This assumption will appear unjustifiable when we take into account Donagan's own discussion of the 'corruption of consciousness'.

2. A principle for determining the scope of prohibitions

The problem of consequentialism only arises in Donagan's argument because he makes a serious attempt to address a crucial issue that other 'deontological' moral theorists (including Kant) have unduly ignored or oversimplified. This is the problem of determining on what grounds the *extent of application* of moral duties – including even absolute prohibitions – should be decided. For Hebrew-Christian 'common morality', the scope of a precept or prohibition tells us what specific actions and circumstances (as they arise) fall under its application. By analogy with jurisprudential reasoning, Donagan suggests that such determinations of scope should be made by a process of 'establishing specificity premises' (p. 71) that determine whether cases of a particular kind do or do not fall under a certain rule. Rather than determining specificity premises for every possible case, however, the moralist should specify the extent of a moral precept in an open-ended limiting procedure that aims to permit as few exceptions to the moral precepts as possible:

A natural approach is to begin by showing that it is impermissible to perform actions of that kind at will, and then to go on to determine

³ However, as Sen points out in his *Inequality Examined* (Cambridge, MA: Harvard University Press, 1992), virtually all moral theories that allow consequences any relevance for interpersonal comparisons start with a particular conception of 'consequences' as measured in terms of some focal variable and space of analysis (see pp. 1–2) and a conception of how they are to be 'combined' in normative analysis (see pp. 73–74). To be rigorous, then, we would have to say that a 'consequentialist' theory in Donagan's sense is one which both (1) admits as relevant only broad external features (such as wealth) or psychological features (such as preference-satisfaction) – rather than kinds of action and intentions, and (2) determines the normativity of rules purely by some logical combination or summation of these sorts of consequences.

the kinds of cases in which it is permissible (p. 72)

By this procedure, we see that certain absolute prohibitions or duties, such as the duty to keep promises, implicitly contain 'tacit conditions' for circumstances in which they will no longer apply.

Although Donagan never explicitly sets out a Kantian principle for deciding whether to make an exception that limits a prohibition, Donagan *implicitly* appeals to the following two-fold principle when he considers how to qualify specific prohibitions and how to defend his Pauline Principle against consequentialist objections:

(Criterion C): To justify adding an escape clause to an absolute prohibition in an extreme kind of circumstance K, we would require almost *certain knowledge* that the results of regularly obeying the moral prohibition in circumstances the same as (or relevantly similar to) K would be *calamitous* in comparison to the consequences of regularly making exceptions in circumstances of kind K.

For such a criterion to make sense, obviously, we must assume that 'calamitousness' can be defined according to some adequate theory of harms.⁴ Ideally, harms (and thus calamitousness) could be defined relative to certain primary goods whose value is obvious and consistent with the fundamental precepts of common morality.⁵

We can see that Donagan appeals to something like principle (C) in his argument for giving a very broad scope to the prohibition against lying. Donagan says that the duty of veracity is 'not unrestricted,' but that even with regard to duping children and the mentally impaired,

... the weight of Jewish and Christian opinion is on the side of veracity, except where it is *beyond doubt* that a truthful statement or evasion will cause *unjustifiable harm* (p. 89 – my italics).

In other words, to justify lying in a particular type of circumstance, we would have to be almost certain that telling the truth in that circumstance would cause terrible results. The determination involves both consideration of our certainty, and consideration of whether immediate disaster is the certain consequence.

⁴ A complete criterion would obviously have to include some way of specifying what can and cannot count as calamitous in the relevant way. However, we can already see that the description of a circumstance allegedly involving risk of calamity must employ only qualitative predicates, excluding properties such as 'having calamitous consequences,' for example. I am indebted for this point to Walter Sinnott-Armstrong.

⁵ Another way to strengthen this criterion would be to include the requirement that the results of obeying the moral prohibition be calamitous in comparison to making a *publicly acknowledged* regular exception for cases of the kind in question. I am indebted to Jerome Schneewind for this point.

Donagan also employs (C) in considering the scope of the precept that 'it is impermissible for anybody within a civil society to disobey its laws' (p. 109). Active disobedience or rebellion is permissible only if the society institutionalizes great wrongs and 'if it can reasonably be predicted that the rebellion will succeed in righting the wrong ...' (p. 109). If we are not absolutely certain that consequences of refraining from rebellion will be more calamitous than the consequences of rebellion itself, this doubt is enough to subsume the case under the prohibition against rebellion *at will*.

Donagan also appeals to (C) in defending several of his moral prohibitions against consequentialist critiques. He suggests that 'the examples commonly offered in support of escape clauses' usually misrepresent the facts by 'imputing to the agents a knowledge they cannot possess' (p. 207). Thus, in the case described by Bernard Williams, where a captain threatens to kill twenty innocent Native Americans unless an unfortunate visiting botanist agrees to execute one of them, Donagan argues that the botanist should 'on no account agree to kill anybody,' partly because he cannot foresee with reasonable certainty that the death of all twenty victims will follow if and only if he refuses to kill one of them as commanded (p. 208). Any doubt about the relative disastrousness of the outcome defeats the warrant for placing this type of case outside the scope of the prohibition.

3. Donagan's insight and inadvertent consequentialism

Donagan's treatment of these kinds of 'hard cases' for common moral precepts involves a misapplication of a more general insight. In trying to determine the scope of moral precepts, Donagan clearly recognizes that a deontological theorist can consider *morally relevant* consequences in deciding the extent of prohibitions, without thereby violating 'deontological' restrictions on the justification of moral principles. By itself, I think this represents an insight of some importance.

A deontological metaethics does require, however, that the moral significance of different kinds of consequences, how they are summed, and the weight or priority some enjoy over others, are all determined by more than allegedly 'natural' differences in the value of consequences themselves. Rather, these determinations reflect critical normative distinctions stemming from the deontological first principle or precepts derived from it. Then the relevant evaluation of consequences can be used in determining the scope of moral precepts without being 'consequentialist' in the strict sense. For example, a prohibition might be limited so as not to include a case in which its consequences would be 'unacceptably disastrous,' if what counts as 'disastrous' were determined on the basis of the first principle of 'common morality' (or some intermediate precept), rather than being

determined solely by some form of summation of all consequences for the well-being of the parties affected.

But in Donagan's treatment of specific cases, unfortunately, what counts as a 'disastrous' outcome seems to depend on standards of evaluation that are consequentialist in the narrow sense. For example, against Williams's dilemma for the botanist, Donagan argues that in addition to the uncertainty involved, the '*acceptance-utility*' of an exception rule for cases of this kind (e.g. the threat of murder or mass killings) probably would not be higher than the acceptance-utility of the prohibition without the exception clause: 'would not general acceptance of the rule that one is to comply with such invitations to murderous complicity be more likely to have bad results than not?' (p. 208). Donagan does not explain this remark in detail, but presumably he means that allowing murders under such duress would only encourage military autocrats to compel complicity from others in the future. Although it is interesting, this argument turns *purely* on a comparison of the long-run combination of consequences for the well-being of the parties involved – and so it is 'consequentialist' in Anscombe's sense.

Likewise, in a case presented by Michael Sandel, where an honest politician must decide whether or not to torture a terrorist to discover where his cadre have planted bombs, Donagan argues that to justify torturing the terrorist, we would first have to know with *certainty* that the terrorist possesses the required information (p. 187). Second, we would have to be certain that a policy of permitting torture in such extreme circumstances is likely to be for the greater good in the long run. Here again, the epistemic and consequentialist criteria in (C) are crucial to Donagan's argument that an exception-rule would not be for the long-term good (p. 188).

By using the two criteria in (C) to judge possible exceptions to moral precepts, Donagan is attempting to show that consequentialist arguments against 'traditional morality' fail by *their own standards*. As appealing as this strategy may be, however, it is perilous. For by basing his arguments on (C) Donagan lays himself open to a host of parallel counter-arguments. First, there is always the possibility that likely consequences will justify **more exceptions than Donagan could accept**. For example, suppose that Williams's botanist has reason to suspect that refusing to make an exception to the prohibition against killing in his circumstance might be disastrous for this society.⁶ If he is certain of this, then (C) immediately justifies murdering the innocent 'scapegoat'. Similar possibilities remain open for Sandel's terrorist case as well.⁷ But since part of Donagan's point is that dilemmas may involve uncertainty, let us imagine that the botanist

⁶ E.g. the botanist knows that in this society, killing one 'scapegoat' out of a condemned group is a deeply embedded social ritual, and that refusing to play his role risks calamitous consequences for the long-term peace of this society.

still has a lot of doubt about the outcome. Then we still have the following problem: just as the considerations introduced in (C) can be used to argue against making an exception to a precept, by parity of reason, it would seem that they would be used to argue *for* making an exception. The consequentialist can argue that since making the exception is not certain to be disastrous, the botanist lacks sufficient grounds for refusing to make the exception and carry out the execution. Unless there is some reason why the presumption must always go *against* the exception clause under uncertainty, a consequentialist can use all of Donagan's own criteria against him.⁸

4. A critique of Donagan's epistemic criterion

Thus Donagan's criterion of calamitousness, at least as he applies it, is not consistent with a deontological approach to determining the scope of moral precepts. I have also suggested that Donagan's second, epistemic requirement of certainty can also be used by the consequentialist to argue for allowing exceptions to moral precepts, thus limiting their scope more narrowly than 'common morality' might expect. However, Donagan does have a reason for assuming that in uncertainty, the presumption must always be against making the exception to a common precept. He suggests that 'since nearly everybody's judgment is disturbed by the anticipation of calamity, it is probable that much of what is done on the grounds of such

⁷ The recognized theory that one should never negotiate with terrorists, despite the cost to their hostages, is based purely on consequentialist considerations of long-run acceptance utility. If these are the sole relevant considerations in this case, they might well be sufficient to justify torturing the terrorist in the Sandel's case.

⁸ In fairness, it should be noted that Donagan avoids deploying (C) in a consequentialist fashion in defending his precept that 'nobody may rightly volunteer to serve in a war unless he is convinced it is a just one' (p. 111). Since a potential recruit must try to ascertain the justice of the war (p. 110), he is prohibited from serving in wars *at will*. For a war to be a just requires that the disaster to be averted by force of arms is great in comparison to the toll of war itself: 'War is so horrible an evil that only a very clear and great cause can justify it...' (p. 111). So for Donagan, one's service in a particular war is an exception to the general prohibition only if one is certain on critical reflection that the probable consequences of not fighting the war would be disastrous enough, as per (C). But prior to that, the war strategy must also respect certain absolute prohibitions, such those against initiating hostilities for territorial gain or employing more force than is needed for defensive purposes (p. 110). Thus there is no pure balancing of consequences in the judgment: the disastrous consequences of *not fighting* the war, and its potential human costs and benefits generally, do not even become relevant to determining the war's justice if its strategy violates more basic prohibitions. Donagan's treatment of this problem illustrates how a deontological theory must condition and qualify any significance it gives to consequences in determining the scope of its precepts.

escape clauses will be mistaken' (p. 207). This 'psychological argument' suggests that the presumption must always be on the side of subsuming the case under the prohibition, because people fearing calamity will be apt to overestimate the real chance of disaster in their circumstances, and thus resort too quickly to any escape clauses that are generally recognized and allowed. In other words, escape clauses should be kept to the barest minimum because they will be overused anyway. Hence we should always take uncertainty about the long-term results as grounds against, and never for, limiting the scope of the precept to exclude a difficult circumstance.

While there is doubtless some merit to this argument, it is too one-sided. If the epistemic criterion in (C) can be used by the 'traditionalist' to argue against making exceptions, the 'realist' in favour of certain escape clauses has just as good a claim to the uncertainty criterion in many cases. The 'realist' can argue that, although people are prone to panic, they are also prone to apathy and prone to seek any justification for *not* 'getting involved.' In extreme circumstances, the danger is often that people will *underestimate* the likelihood that calamity will result from failing to make exceptions to traditional precepts.⁹ Especially in difficult circumstances where making the exception might require decisive action, risk, or sacrifice on their part, people are more apt to stick to 'tried and true' precepts, rather than to make an exception when they should. In these circumstances, where there is already a base motive for acting in accordance with the precept, general public recognition for an escape rule limiting a traditional prohibition might encourage precisely the kind of critical reflection that can protect people from the 'corruption of consciousness.'

In his discussion of the 'corruption of consciousness,' Donagan explains that people tend to avert their attention from what 'disturbs' them (p. 141). But, if people's judgment is also 'disturbed by the anticipation of calamity,' as Donagan maintained in his 'psychological argument,' then surely they may tend to avert their attention from indications that calamity is *likely*. If evidence of the risk of calamity may disturb a person's judgment, out of a desire to reduce his or her anxiety, a person may choose to ignore or undervalue such evidence.¹⁰

⁹ Consider, for example, the precept not to intervene in another nation's internal affairs. Perhaps no other precept has been *overextended* in its scope. I take the prolonged, agonizing refusal of Western powers (especially the European Community) to act decisively against Serbian fascism to be a case in point: for the sake of inaction, popular opinion refused to face the obvious fact that the incentive for further ethnic aggression and genocide spreads throughout the world as a result.

¹⁰ The psychological theory of cognitive dissonance, developed by Leon Festinger in the 1950s, supports this hypothesis. See Festinger, *A Theory of Cognitive Dissonance* (Stanford, CA: Stanford University Press, 1957).

Therefore Donagan's own psychological argument implies that there can be a close connection between extreme circumstances, fear of calamity, and corruption of consciousness. Presumably it is especially disturbing to conscientious people to think that obeying a traditional moral precept to the letter could lead to a calamitous consequence: persons with a strong disposition to follow moral prohibitions taught by their community may therefore be particularly susceptible to underestimating the extent to which, in extreme circumstances, normally appropriate responses may involve risk of calamity. This is especially true if the calamity at risk is not one that would affect their own well-being directly, while they would have to put themselves at risk to depart from their accustomed practices.¹¹

Suppose, for instance, that a conscientious and patriotic citizen has reluctantly agreed to hide members of a threatened ethnic minority in her home. Now she is asked by an official whether she is harbouring anyone. Let us suppose that she does not know *for certain* that they will be mistreated if she gives them up: she has to interpret ambiguous evidence from widespread rumours that she has heard on the streets and what little information she has gathered at her job.

Here our concern ought not to be that she will overestimate the danger of calamity to her protectees, as Donagan's reasoning would seem to imply, but precisely the reverse. If she can convince herself that her protectees' peril is not all that great, she can avoid the risks of retribution from the authorities. It will go much easier for her if she can rationalize that the rumours she has heard are probably all spurious: thus she avoids breaking the prohibition against lying when the official appears at her door. She desires, after all, to behave as she is used to doing, obeying traditional laws and precepts to the letter, being a truthful person, respecting officials, and believing that all the troubles will pass. She cannot even allow herself to suspect that things have gotten so out of hand that even the basic duty to be honest is no longer appropriate.

Moreover, this woman's self-deceiving self-assurance illustrates a danger that extends beyond individuals and isolated incidents. The 'corruption of consciousness,' as Donagan admits, may become so inveterate that even 'a whole society may refuse to face plain facts' (p. 140). If this is true, then

¹¹ Donagan himself acknowledges something close to this possibility in discussing the heroism of Franz Jägerstätter in refusing to honour the Nazi draft, despite *opposition* from his bishop (pp. 16–17). Donagan explains the poor clerical advice Jägerstätter received by attributing it to the 'depravation of the *Sittlichkeit*' of their community. But perhaps the bishop's unwillingness to make an appropriate exception to the first-order prohibition against disobeying 'lawful civil authority' induced him to repress the evidence of the war's injustice, allowing him to regard the case as a 'doubtful' one which would thus be covered by the prohibition (as per p. 111)?

whole segments of a society may deceive themselves about the probable calamity of not making an appropriate exception to a given moral precept in an extreme circumstance, even when the warrant for it is obvious.

For example, assuming that there is a moral prohibition against brinkmanship in international relations, consider a peaceful nation N faced with an increasingly menacing and corrupt foreign regime. N's leaders already have a motive not to violate the prohibition against brinkmanship in this circumstance, since challenging the corrupt regime is risky for them and politically unpopular. In this circumstance, it is easy for them to refuse to see that calamity will result unless they begin strong initiatives to oppose the corrupt regime. In false consciousness, they may be unable to believe the mounting evidence that the danger of not preparing is worse than the danger of provoking the potential enemy by a hostile buildup of their own forces.¹²

It follows that in circumstances where the corruption of consciousness is a real danger because there are untrustworthy motivations *not* to make an exception to a moral precept, Donagan's uncertainty criterion would be better used to support making the exception and limiting the precept.

In conclusion, both of Donagan's criteria for determining the scope of moral precepts turn out to be inadequate for a deontological theory. Deciding the limits of a precept's application purely by a comparison of acceptance-utilities is incompatible with a deontological system, and it is more likely than Donagan admits to justify exceptions to prohibitions in 'hard cases'.¹³ Moreover, his epistemic criterion has more ambiguous implications for the scope of moral precepts than Donagan realizes.¹³

University of Notre Dame
Notre Dame, IN 46556, USA
JDavenport@darwin.cc.nd.edu

¹² This case is obviously similar, in some respects at least, to the history of the 'appeasement' of Hitler leading up to the Second World War.

¹³ An earlier version of this paper was presented at a conference on *The Moral Philosophy of Alvin Donagan*, organized by Alasdair MacIntyre and held at the University of Notre Dame (Notre Dame, Indiana, April 14-17, 1994). I wish to express my thanks to Professor MacIntyre for his encouragement, to the other participants at the conference for their suggestions, and to an anonymous reviewer at *Analysis* for criticisms that resulted in substantial improvement of this paper.

Gay rights and affirmative action: a response to Sartorelli

MICHAEL RUSE

In my book, *Homosexuality: A Philosophical Analysis*, having expressed a general distaste for non-job-related preference policies in hiring, for instance through the use of quotas, I argued that there are special reasons why such proposals are inappropriate in the case of homosexuality. Joseph Sartorelli criticizes my position as showing 'a considerably less than complete sensitivity to the plight of gay people in the context of present day Western societies' ([3] p. 84). This may be so; but I would argue that none of the three objections raised by Sartorelli are compelling. Indeed they point to conclusions I endorse.

My first argument centred on the matter of definition. I argued that any definition of 'homosexual' would be so problematic that one could not hope to mount a coherent argument in favour of affirmative action for homosexuals. In this, the situation is other than that between men and women where the differences are clear cut. Sartorelli replies that we have trouble defining 'Blacks' and 'Hispanics', yet this does not stop the passing and enforcing of laws mandating such action for them. 'Unless there is some reason why the existence of unclear cases should prevent us from applying the system in the clear cases, we should not rule out having such a system at all' ([3], p. 85).

I cannot see that this is an adequate counter to my position. Apart from anything else, one surely needs something a little stronger than Sartorelli offers if we are to have a functioning system. Unless the clear cases outnumber – I would think, much outnumber – the unclear cases, one is going to have an unworkable situation. If one could tell only in one case in a hundred if someone were unambiguously black or not, one could hardly have quotas for blacks. And if you tell me that such ambiguity is truly the case now for blacks (even though it is not), then I reply that this is one more argument **against reverse discrimination in the case of blacks rather than one more argument for such discrimination in the case of homosexuals.**

My main point – one which I make at great length in my book – is unchanged. You run into horrendous conceptual problems when you try to define 'homosexual' in such a way as could be used in discrimination laws. Is one to consider only behaviour or feelings (orientation) or both or neither? Should the man just released from prison, with ten years of steady homosexual activity to his record, be considered a homosexual, even if his erotic interests are exclusively in females? Should the celibate nun with a