

Against Maximizing Act-Consequentialism (April 6, 2013)

in *Moral Theories* edited by Jamie Dreier (Blackwell Publishers, 2006), pp. 21-37

1. Introduction

Maximizing act consequentialism holds that actions are morally permissible if and only if they maximize the value of consequences—if and only if, that is, no alternative action in the given choice situation has more valuable consequences.¹ It is subject to two main objections. One is that it fails to recognize that morality imposes certain constraints on how we may promote value. Maximizing act consequentialism fails to recognize, I shall argue, that the ends do not always justify the means. Actions with maximally valuable consequences are not always permissible. The second main objection to maximizing act consequentialism is that it mistakenly holds that morality requires us to *maximize* value. Morality, I shall argue, only requires that we satisfy (promote sufficiently) value, and thus leaves us a greater range of options than maximizing act consequentialism recognizes.

The issues discussed are, of course, highly complex, and space limitations prevent me from addressing them fully. Thus, the argument presented should be understood merely as the outline of an argument.

2. What is consequentialism?

Act utilitarianism is the paradigm *act-consequentialist* theory. It holds that an action is permissible if and only if the aggregate (e.g., total or average) well-being that it produces is no less than that produced by any alternative feasible action. Rule utilitarianism is the paradigm *rule-consequentialist* theory. It holds, roughly, that an action is permissible if and only if conforms to rules that, if generally followed (internalized, upheld, etc.), would produce

aggregate well-being that is no less than that produced by any feasible alternative set of rules. Rule utilitarianism does not assess actions on the basis of the value of their consequences. Instead, it assesses them on the basis of their compliance with selected rules, and it selects rules on the basis of the value of the consequences of their being generally followed (etc.). In what follows, we shall restrict our attention to act-consequentialism, and “consequentialism” and its variants should be so understood. It is beyond the scope of this chapter to assess rule consequentialism.²

Consequentialism has been defined in several distinct ways, but all agree that it involves at least the following two claims:

Supervenience (of Permissibility on the Value of Outcomes): The permissibility of actions in a given choice situation supervenes on (is fully determined by) the value of their consequences.

Value promotion: If, in a given choice situation, one action is permissible, and a second is more valuable, then the second action is also permissible.

Supervenience claims that the moral permissibility of actions in choice situations is fully determined by the value of their consequences. In a given choice situation, two actions that have equally valuable consequences also have the same permissibility status (either both are permissible or neither is). Moreover, if two choice situations (of two different agents or of one agent on different occasions) are *value-isomorphic*—that is, have the same number of feasible actions and are such that their actions can be paired so that the value of the consequences of a given action in one choice situation is the same as its counterpart in the other choice situation—then the choice situations are *permissibility-isomorphic*—that is, an action is permissible if and only if

its counterpart also is. This leaves open how permissibility is determined by the value of consequences. For example, a theory that holds that an action is permissible if and only if it *minimizes* the value of consequences satisfies Supervenience. Value Promotion adds that the supervenience must take the form of value promotion: if, in a given choice situation, one action is permissible, then so is any action with equal or greater value. Value maximization is one way of satisfying this condition, but satisfication (i.e., promotion to some non-maximal adequate level) does so as well.³

Call a theory that satisfies Supervenience and Value Promotion *core consequentialist*. Such theories can involve any account of the relevant value (or values) that is (are) to be promoted. The value might, for example, be (agent-relative) prudential value (roughly: quality of life or well-being) for the agent only (e.g., as with ethical egoism), prudential value from some agent-neutral perspective (taking everyone's prudential value into account; e.g., utilitarianism), moral value from some agent-neutral perspective (e.g., the extent to which prudential benefits match desert), or (agent-relative) moral value from the perspective of the agent (e.g., that, from the perspective of the agent, views it morally worse, for her to kill someone than for someone else to kill that person).⁴ (Many authors call a theory "consequentialist" only if it has an agent-neutral theory of value, so it's worth keeping in mind that I use the term more broadly.) It also leaves open what the correct conceptions are of the relevant kinds of value. It leaves open, for example, (1) whether prudential value is based on happiness, preference-satisfaction, or some perfectionist notion of flourishing, and (2) whether moral value is based on prudential value and, if so, how (total, average, leximin, degree of equality, extent to which it matches desert, etc.).

Core consequentialism can be understood broadly to cover theories that recognize several relevant values (e.g., the prudential value for each agent) rather than simply one overriding value. For simplicity, however, I shall restrict my attention to core consequentialism in the

narrow sense, which requires that permissibility supervene on a single value.

“Consequentialism” should thus be understood as “single value consequentialism”.

I shall object to both (1) the failure of the maximizing consequentialism’s failure to recognize that morality typically leaves us a significant array of *moral options* (morally permissible choices in given choice situation), and (2) core consequentialism’s failure to recognize that there are *constraints* that limit how value may be permissibly promoted. The first objection applies only to maximizing consequentialist theories. The second applies to all core consequentialist theories.

These objections are not new. I shall simply be developing well known points. Furthermore, these objections rest on the correctness of certain views of common sense morality, which defenders of act consequentialism have questioned (roughly by arguing that moral options and moral constraints are irrational).⁵ Although I shall answer some these challenges in passing, I shall not give a full rebuttal. Thus, the objections here raised to act-consequentialism are merely presumptive at best.

3. Against Maximizing Consequentialism

Standard utilitarianism is a paradigm example of maximizing consequentialism. It holds that an action is permissible if and only if the agent-neutral value of its consequences is no less than that of any feasible alternative. One of the main objections to it—and, we shall see, all core consequentialist theories—is that it fails to recognize any (deontological) constraints on the promotion of value. *Constraints* (also called “restrictions”)—such as a constraint against killing innocent persons—rule actions impermissible independently of the value of their consequences. I shall develop this objection in the next section. For the purposes of this section, however, we shall set this issue aside. Our topic here is whether morality requires that, perhaps subject to

some constraints, we maximize value. I shall argue that it does not.

The main argument for the impermissibility of actions that do not maximize value is the following (see, for example, Kagan 1989):

P1: An action is morally permissible only if it is best supported by moral reasons for action.

P2: The value of consequences is always a moral reason for action.

P3: The value of consequences is the *only* moral reason for action—except perhaps for certain prior constraints.

C: Thus, an action is morally permissible only if, perhaps subject to certain prior constraints, it maximizes the value of consequences.

To adequately assess this argument, we need to distinguish between insistent and non-insistent moral reasons for action. *Insistent* moral reasons for action (in a particular choice situation) are considerations that count for, or against, the moral *permissibility* of actions, whereas *non-insistent* moral reasons are considerations that *merely* count for, or against, the relative moral *desirability* of actions (but not for their moral permissibility). Thus, for example, according to common sense, although, under normal conditions, there is a non-insistent reason for me to carefully wrap my wife's birthday present (since it makes her happier), that reason is not an insistent reason. It is morally desirable for me to wrap it carefully, but that consideration is not relevant for the determination of whether it is permissible for me to do so (or not do so).

If P1 is to be plausible, moral reasons must be understood as insistent reasons. By definition, non-insistent reasons (in a given choice situation) are irrelevant to moral permissibility (in that situation). In what follows, then, we shall understand the argument as invoking insistent moral reasons only and not moral reasons generally.

Consider first P3, which claims that the value of consequences is the *only* moral reason for action—except perhaps for certain prior constraints. Given that it allows that there may be constraints on the promotion of value, it is, I believe, highly plausible. Some might wish to challenge it on the grounds, for example, that the intentions of agents provide insistent moral reasons for actions (in addition, or instead of) the consequences of their actions. I, however, am happy to accept P3.

Consider now P2. We can grant that the value of consequences is always a reason for action. P2, as we are now interpreting it, however, claims something much stronger. It claims that the value of consequences is always an *insistent* moral reason for action. Some radical deontologists might deny it is *ever* such a reason, but almost everyone else accepts that it often is. The more controversial part is the claim that the value of consequences *always* is an insistent moral reason for action. Many would argue that, above some (perhaps context-specific) non-maximal but adequate level, the value of consequences ceases to be an insistent moral reason. Consequences must be good enough, but, beyond that, their value is not relevant to determining the permissibility of actions.⁶ This, then, is one objection to the argument.

Consider finally P1. It claims that an action is morally permissible only if it is best supported by insistent moral reasons. It is not obvious, however, this is so. It may be that permissibility only requires *adequate* (as opposed to best) support. To this, it may be objected that this makes morality irrational. Rationality, it may be claimed, requires that only the best supported actions be judged permissible. In reply to this objection, we can note, first, that it is not a conceptual truth that moral permissibility is sensitive to moral reasons in the same way that prudence is sensitive to prudential reasons. Thus, even if best support is required by prudential reason, it may not be required by moral reason. Furthermore, it's not even obvious that reason in general, or prudential rationality in particular, requires best support (by insistent reasons). The

satisficing model of rationality has been defended by many, and it only requires adequate support for rational permissibility.⁷

Obviously, the issue is complex. My claim is that the above argument for (perhaps constrained) maximization of consequentialist value is not valid as a conceptual matter (given the concepts of moral permissibility, reasons for actions, etc.). This leaves open, of course, that (constrained or unconstrained) maximizing consequentialism may nonetheless be true. In the remainder of this section, I shall identify three substantive problems that maximizing consequentialist theories confront. If my analysis is correct, then we have good reason to reject maximizing consequentialism.

Standard utilitarianism is subject to the following three closely related objections: (1) it often requires agents to sacrifice excessively their well-being, (2) it leaves agents inadequate moral freedom (judges too few actions morally permissible), and (3) it leaves no room for permissible actions that are morally better than other permitted actions. I shall explain each below and show how, with certain qualifications, these objections apply to maximizing consequentialist theories generally.

One problem with utilitarianism is that it requires agents to make significant sacrifices in their own well-being in the name of value maximization. Only those actions that maximize value are judged permissible, and these often require agents to make significant sacrifices in their own well-being. The objection here is *not* that utilitarianism *sometimes* requires agent's to make significant sacrifices; all plausible moral theories have this feature. Any plausible theory, for example, will typically judge it impermissible to steal a million dollars, even though one can get away with it and would greatly benefit from the result. The objection here is that utilitarianism *frequently* requires significant sacrifices from agents. It holds that typically it is wrong to spend money (e.g., for restaurants, clothes, or CDs) or time (e.g., watching TV, talking with friends) for

one's own enjoyment, since usually this money or time would provide a greater benefit to disadvantaged persons. Of course, such activities are not always wrong, since the most effective way of helping others typically involves occasionally pampering oneself. Most of the time, however, utilitarianism judges it impermissible to devote more than minimal time or resources to oneself.

One way of countering, or at least weakening, this objection is to insist that, when assessing the sacrifices demanded by morality, one must factor in the *passive benefits* that morality provides by requiring *others* to make sacrifices for one's benefit. One must not, that is, only look at the cost when morality requires one to sacrifice one's well-being for the benefit of others, one must also look at the benefit when morality requires others to sacrifice their well-being for one's benefit.⁸ This is an important issue that I cannot here address adequately. The main point on which I would insist is that there are significant limits on the extent to which morality requires one to actively sacrifice one's well-being. Active and passive benefits are not interchangeable. The mere fact, for example, that the level of required sacrifice is less than the passive benefits that morality provides (so that one comes out ahead even with the sacrifice) is not sufficient to establish that morality requires such sacrifice. The issue, however, is complex, and in what follows I shall simply assume, without adequate argument, that this is so.

The problem of excessive required sacrifice is, I believe, a powerful objection to utilitarianism. It is not, however, applicable to all maximizing consequentialist theories. Ethical egoism—which judges an action permissible if and only if it maximizes the agent's well-being—does not require excessive sacrifice. Nor do theories that give great weight to the agent's well-being as compared with others: for example, a theory that gives 50% of the total weight to the agent's well-being (and the other 50% spread equally among the well-being of others). Thus, if the value that is to be maximized is agent-relative and sufficiently *agent-favoring*, this objection

does not apply to maximizing consequentialist theories.⁹

Even if the maximized value is *agent-neutral* (i.e., gives no special role to the agent), certain kinds of maximizing consequentialist theories are immune to the problem of excessive required sacrifice. It depends on whether the theory of value is relatively fine-grained and complete. A theory of value is *relatively fine-grained* just in case it makes sufficiently many distinctions in value so that ties are relatively unlikely. A theory of value is *relatively complete* just in case incomparability of value (neither of two outcomes being at least as good as the other) is relatively unlikely. Utilitarianism, at least its standard versions, has a fully fine-grained and complete theory of value. Thus, few feasible actions maximize agent-neutral value, and thus few are judged permissible. Given that value is agent-neutral, this means that agents will often be required to make significant sacrifices. If, however, value is relatively coarse-grained or incomplete, there may be many actions that maximize value. Suppose, for example, that value is coarse-grained in that outcomes have only three possible values—good for outcomes that are significantly better than the status quo, bad for outcomes that are significantly worse than the status quo, and neutral for all others. Typically, there will be many actions that maximize value. Likewise, if there is a lot of incomparability in value (e.g., incomparable whenever it is better for some people and worse for others), then there will many actions that are judged permissible (since many actions will be such that no alternative action has more value). As long as the theory of value is not agent-disfavoring (i.e., is agent-neutral or agent favoring), typically at least one of the permissible actions will not involve significant sacrifice.¹⁰ Thus, if value (even if agent-neutral) is relatively coarse-grained or incomplete, the problem of excessive required sacrifice need not arise.

My first objection is thus that maximizing consequentialist theories based on value that is agent-neutral (or agent-disfavoring) and relatively fine-grained and complete require excessive

sacrifices of agents. This objection does not apply to theories for which the value is (agent-relative and) sufficiently agent-favoring, nor to theories that have relatively coarse-grained or incomplete values.

A second objection to standard utilitarianism is that it leaves agents too little moral freedom. *Moral freedom* is the extent to which morality leaves agents free to choose among their feasible actions. Of course, morality does not leave agents perfectly free. It judges many actions impermissible (such as killing or assaulting innocents under normal conditions). The issue here concerns how much morality limits our freedom to choose. Utilitarianism effectively eliminates this freedom (except for the case of the occasional tie for best consequences), whereas common-sense, which I claim is correct on this topic, holds that morality leaves us free to choose among a fairly significant range of choices (and thus leaves agents lots of room to decide how to live their lives).

The problem of limited moral freedom is distinct from that of excessive required sacrifice. Ethical egoism is not subject to the problem of excessive required sacrifices, but, if it has a relatively fine-grained and complete theory of value, it leaves agents little moral freedom. Only maximally good options are judged permissible, and almost no option is maximally good. Thus, the problem of insufficient moral freedom can arise when there is no problem of excessive required sacrifice. Furthermore, the problem of excessive required sacrifice can arise when there is no problem of insufficient moral freedom. Consider the bizarre and wildly implausible theory of “satisficing anti-egoism”, which holds that an action is permissible if and only if it is among the worst 30% for the agent. This leaves the agent a fair amount of moral freedom, but still requires excessive sacrifice from the agent.

The problem of limited moral freedom applies to all maximizing consequentialist theories with relatively fine-grained and complete theories of value—even if they are not agent-neutral.

My claim is that morality recognizes a significant range of *moral options* (or *prerogatives*), which are morally optional actions (that is, actions that are permissible but not morally required). If, in a given choice situation, only one action is permissible, then there are no moral options; the one action is obligatory. If, however, more than one action is permissible, then each permissible action is optional. The greater the number of actions that are permissible, in a given choice situation, the greater the extent to which morality recognizes moral options (and hence moral freedom).¹¹

So far, then, we have two independent (but related) objections to maximizing consequentialist theories with relatively fine-grained and complete theories of value: (1) when based on agent-neutral value, they require excessive sacrifices from agents, and (2) they leave agents little moral freedom. The first objection is concerned with the impact on the agent's well-being, whereas the second is concerned with the moral freedom left to the agent's will.¹²

A third problem with utilitarianism is that it holds that no actions are supererogatory. *Supererogatory actions* are actions that are morally optional (permissible but not obligatory), are significantly more valuable than other permissible alternatives, and involve significant sacrifice by the agent. Thus, for example, according to common-sense, devoting all one's income and time to help others is supererogatory. Utilitarianism does not recognize supererogatory actions because it holds that no permissible action is more valuable than any other permissible action (since only maximally valuable actions are permissible). Common-sense, which I claim is correct in this respect, however, holds that some permissible actions are better than others. It may be better for me to help a poor fatherless child, but I am permitted to watch television instead. Such help is supererogatory, not obligatory.

The problem of insufficient supererogation is closely related to, but distinct from, the problem of the lack of moral freedom identified above. It is closely related in that if moral

freedom is severely limited, then so is the range of supererogatory actions (since an action is supererogatory only if it is optional). The problem is distinct, however, since abundant moral freedom does not entail any supererogation. A maximizing theory with a relatively coarse-grained or incomplete theory of value will leave a significant amount of moral freedom, but it does not recognize that some permissible actions are better than others (they are all either equally good or incomparable).

The problem of insufficient supererogation is also distinct from the problem of excessive required sacrifice. Satisficing anti-egoism (e.g., an action is permissible if and only if its outcome is among the worst 30% for the agent) recognizes sufficient supererogation but imposes excessive required sacrifice. Furthermore, ethical egoism fails to recognize supererogation but involves no excessive required sacrifice.

With one qualification, the problem of insufficient supererogation applies to all value maximizing theories. It applies even to theories for which value is relatively coarse-grained or incomplete, since even such theories judge an action permissible only if no action has better consequences. The qualification is that the value that is used to assess betterness for supererogation must be the same value that is used to assess betterness for permissibility. If these are different values (e.g., the value of consequences is used for permissibility and the value of the agent's intentions is used to assess betterness for supererogation), then, of course, not all permissible actions—those that maximize one value—need be equally valuable according to the second value. Such a detachment of the value for permissibility from that for supererogation is, however, highly implausible, and I shall ignore this possibility.¹³

The three problems raised have, as I have indicated, different scopes. The problem of insufficient supererogation applies (with the above qualification) to all maximizing consequentialist theories. The problem of insufficient freedom applies to all such theories with

relatively fine-grained and complete theories of value. The problem of excessive required sacrifice applies to all such theories with agent-neutral theories of value (or more generally: insufficiently agent-favoring theories of value).¹⁴

It's worth noting that these three objections arise independently of whether the value of an action is determined by the value of its consequences. They arise for any maximizing theory, no matter what the basis is for evaluating actions, as long as value is relatively fine-grained and complete, and, for the problem of excessive sacrifice, agent-neutral. Suppose, for example, that the value of actions has these features, is determined by the underlying intentions, and morality requires that agents maximize the value of actions. Such a theory requires agents to sacrifice excessively their well-being, leaves them inadequate moral freedom, and does not recognize any supererogatory actions.

If, as I believe, a plausible theory must avoid these three problems, then agent-neutral consequentialist theories must be satisficing rather than maximizing. There are many forms that satisfaction can take depending on how the satisfactory level of value is determined: for example, some percentage (e.g., 80%) of the maximum feasible value, some percentile value (e.g., the option that is at least as valuable as 80% of the feasible options), or the value produced if one maximizes a weighted function that gives extra weight to the agent's well-being (roughly the view defended by Scheffler (1982)). There are, of course, many other possibilities, and I shall not here attempt to specify any particular version.¹⁵

4. Against Core Consequentialism

Core consequentialism holds that the permissibility of actions supervenes on (is fully determined by), and is positively sensitive to, the value of their consequences. This does not require value-maximization; it only requires value promotion, where the relevant value is that of the

consequences. I shall argue that core consequentialism, at least in its standard forms, is mistaken. The ends do not always justify the means.

If core consequentialism is true, then any action with maximally good consequences (in a given choice situation) is permissible. The main argument in favor of this claim is the following:

P1: An action is morally permissible if it is best supported by insistent moral reasons for action.

P2: The value of consequences is always an insistent moral reason for action.

P3: The value of consequences is the *only* insistent moral reason for action.

C: Thus, an action is morally permissible if it maximizes the value of consequences.

This is the same argument given in the previous section for the impermissibility of actions that do *not* have maximally good consequences, except that (1) the appeal to *insistent reasons* has been made explicit, (2) the necessary conditions of the original P1 and C have been converted to sufficient conditions, and (3) the qualification in P3 that allowed the possibility of some prior constraints has been dropped.

P1 is highly plausible. An action that is best supported by insistent moral reasons is surely permissible. P2 can be challenged, as I did earlier, on the ground that beyond some point the value of consequences ceases to be an insistent moral reason (once consequences are good enough, their value may only be a non-insistent reason). For the present purposes, however, we can grant this claim. The crucial claim is P3. It is implausible, because there are insistent moral reasons other than the value of consequences. There are also deontological insistent reasons, and these, or at least some of these, are lexical prior to the value of consequences. In particular, individuals have certain rights that may not be infringed simply because the consequences are better. Unlike prudential rationality, morality involves many distinct centers of will (choice) or

interests, and these cannot simply be lumped together and traded off against each other.¹⁶

The basic problem with standard versions of core consequentialism is that they fail to recognize adequately the *normative separateness of persons*. Psychological autonomous beings (as well, perhaps, as other beings with moral standing) are not merely means for the promotion of value. They must be *respected* and *honored*, and this means that at least sometimes certain things may not be done to them, even though this promotes value overall. An innocent person may not be killed against her will, for example, in order to make a million happy people slightly happier. This would be sacrificing her for the benefit of others.

The claim here is that there are some constraints on how value may be promoted. The ends do not always justify the means. Moreover, these constraints, as I shall explain below, are grounded in the normative separateness of persons.¹⁷

Constraints may be personal or impersonal. An *impersonal constraint* against killing, for example, prohibits killing, independently of whether this is in the killed person's *interests* and independently of whether she has *consented to it* (i.e., is in conformance with her will). It would rule out, for example, well-informed suicide, voluntary euthanasia, and non-voluntary euthanasia where an incompetent individual is terminally ill and likely to be in great pain for the remainder of her life. Although impersonal constraints do reflect a normative separateness of individuals, they do not do so, I believe, in the relevant manner. They fail to capture the *respect due to persons*. Persons (beings that are protected by morality for their own sake) have interests and often autonomous wills. Any constraint against treating a person in a specified way that applies even when the holder validly consents to such treatment and such treatment is in the holder's interest fails to reflect the respect due to that person. Impersonal constraints fail to reflect this respect, and I agree with core consequentialism's rejection of such constraints.

Constraints can, however, be *personal*. A personal constraint empowers the protected

individual, and makes the prohibition conditional on it thwarting her interests or, alternatively, not being in conformance with her will. Personal constraints are waivable rights, and are waived (and hence not violated) when the breach of the constraint is—for *interest-protecting rights*—in the person’s interests, or—for *choice-protecting rights*—when the person has given valid (e.g., free and informed) consent.¹⁸ Thus, for example, well informed suicide and voluntary euthanasia do not violate the choice-protecting right against being killed, and non-voluntary euthanasia for a person with a life not worth living does not violate her interest-protecting right against being killed.

Personal constraints—both choice-protecting and interest-protecting rights—reflect the normative separateness of persons in an appropriate manner. Like impersonal constraints, they require that the holder not be used merely a means for promoting value. Unlike impersonal constraints, by giving a special role to the interests or will of the rights-holder, they further require that the holder be treated with respect.¹⁹

There are, of course, many important questions that need to be answered. One concerns the content of the rights. For the present purposes, we don’t need to answer this question. All that matters is that there are some rights. I believe, for example, that one of our core rights is that of bodily security (e.g., against being killed, struck, or restrained). A second issue concerns whether the rights are choice-protecting or interest-protecting. The issue concerns the nature of the requisite respect that rights require. I’m inclined to think that psychologically autonomous agents have (mainly) choice-protecting rights and non-autonomous but sentient beings (such as young children and certain animals) have interest-protecting rights, but we need not resolve this issue here. All we need is the existence of some kind of right.

A third issue concerns whether the rights are absolute or conditional in certain ways. Rights with thresholds, for example, have no force when the value that would be foregone is

above some threshold (e.g., a right against being killed might not apply where infringement is the only way of avoiding social catastrophe).²⁰ If there are thresholds, then at some point the normative separateness of persons yields to the promotion of value. For the present purposes, we can leave this open. As long as the rights at least sometimes have some force, the normative separateness of persons will be at least partially recognized in a way incompatible with core consequentialism.

The objection to core consequentialism is that it does not recognize that the ends do not always justify the means, and more specifically that the normative separateness of persons (as reflected in rights) make it impermissible to treat people in certain ways even if it promotes value.

All standard forms of core consequentialism are subject to this objection, but it's worth noting the certain non-standards forms are immune to it. If the value that agents are required to promote is sensitive in a certain way to the violation of the relevant constraints, the objection does not apply. For simplicity, I shall assume that the relevant constraints are absolute (have no thresholds), but the same point can be made (more clumsily) without this assumption. Suppose, first, that violation of the relevant constraint is a source of disvalue, and that such disvalue is *lexically priori* in importance to any other sources of value. Thus, an outcome in which a constraint is violated is always worse (agent-neutrally) than an outcome in which the constraint is not violated—even if large numbers of people are much better off when the constraint is violated. A theory like this might hold that it is better to have one violation now rather than five later (this is sometimes called “utilitarianism [more strictly: consequentialism] of rights”). To block this, suppose further that the theory of value holds that there is a lexical priority of importance to *earlier times* with respect to constraint violation. Thus, an outcome in which the earliest time at which a constraint is violated is t is always worse than one in which the

earliest constraint violation is later than t . With this condition built in, where it is possible to perform an action that does not violate a constraint, a maximizing consequentialist theory with this theory of value will judge it wrong to violate a constraint.²¹

Such a theory is not, however, totally immune to the objection. Where it is impossible to perform an action that does not violate a constraint, there is a difference between such a theory and a standard kind of constraint theory. A maximizing theory will always judge some constraint-violating action permissible, since there is always some action the consequences of which are no worse than those of any alternative action. It would hold it permissible, and perhaps obligatory, to infringe now one person's rights where the only alternative is to infringe now the same rights of two others. A constraint theory, however, will judge all feasible actions impermissible, if its constraints apply even in the case where it is impossible to avoid violating a constraint. (Some constraint theories, of course, may hold that the constraints do not apply in such cases.) There is thus an important sense in which such a maximizing consequentialist theory might still fail to recognize the normative separateness of persons. For the present purposes, however, we shall ignore this relatively small difference. A maximizing consequentialist theory based on the above constraint-violation theory of value will be at least roughly equivalent—with respect to permissible constraint violation—as a theory that imposes genuine constraints.

My objection against core consequentialism, then, does not apply to maximizing consequentialist theories with the above theory of value (and related theories). In what follows, I limit my criticism to standard kinds of consequentialism.

It's worth noting that the objection *does* apply to a certain kind of theory that has sometimes been thought to immune to the objection. Suppose that value is *agent-relative*, and that constraint violation by the agent (as opposed to others) is lexically prior to other considerations of value—but *with no temporal lexical priority* concerning constraint violation.²²

Thus, from the agent's perspective, her violation of any constraint is always worse than any outcome in which she does not violate any constraints (even if others violate many constraints). This agent-relative concern for the agent's own violations is not sufficient to avoid the problem of failing to respect the normative separateness of persons. Such a theory would judge it permissible, indeed, obligatory, to violate a constraint now, where the only alternative is not violating any rights now, but then violating several later. Such a theory is concerned with "keeping the agent's hands as clean as possible" in a consequentialist way, but it does not adequately respect the normative separateness of persons. It judges it permissible for Jones to kill Smith in order to avoid Jones' later killing four other people.²³ This fails to respect the normative separateness of Smith from the others.

Before concluding, we need to consider the main objection to constraints. They are, it is claimed, irrational. There are several different versions of this claim. One is that it is simply irrational to deem the action with the best consequence impermissible merely because it infringes a constraint. This is indeed irrational, if the only insistent moral reasons are the values of consequences. We have seen, however, that there is good reason to deny this claim (since there are also deontological reasons protecting the normative separateness of persons). A second version of the irrationality claim is that, if constraint infringement is a source of disvalue, then surely, all else being equal, it should be permissible to infringe a given constraint when this is necessary and sufficient to minimize the infringements of that constraint. Surely, it is permissible to murder one innocent person when this is the only way to prevent four innocent persons from being murdered. This would indeed be permissible if constraint infringement were simply a source of disvalue. We have seen, however, that constraints do much more than this: they recognize the normative separateness of persons. Thus, the infringement of a constraint protecting one person cannot be simply traded off against infringements of that constraint (or

other ones) protecting others. What matters is the respect of persons—not the respect of the constraints.

A third version of the irrationality claim has more force. It acknowledges the normative separateness of persons, but claims that it is irrational to judge it impermissible to infringe a constraint protecting a person when this minimizes the violations of constraints *protecting that very same person*. There are different versions of this objection, but at the core is the relatively weak claim that it is irrational to judge it impermissible for a given agent to infringe a given constraint protecting a given person when this minimizes the number of infringements by the *same agent* of the *same constraint* protecting the *same person*. It is irrational, for example to judge it impermissible for Smith to steal \$100 from Jones now when this minimizes the number of times that Smith steals \$100 from Jones. This weak form of the objection is silent when different constraints or different agents are involved, but the objection still has a significant amount of force when different constraints and different agents are involved. The crucial point here is that, in this kind of case, the tradeoffs in constraint infringement are all *intrapersonal*—all involve the same protected person—and not interpersonal.

These are important points, and they may well show that it is irrational, or otherwise implausible, to prohibit constraint infringement when it minimizes constraint infringements in these victim-focused ways.²⁴ The central issue to be resolved is, I believe, what kind of respect is owed to persons and whether infringement-minimizing infringements are compatible with such respect. I shall not, however, here attempt to resolve this issue.²⁵ The important point is that, even if infringement-minimizing infringements are permissible, this does not show that there are no constraints on the promotion of the value. It would only show that the constraints are conditional and do not prohibit the specified actions when those actions have the requisite constraint-infringement minimizing feature. There would still be applicable constraints on the

specified actions that do not minimize constraint violation in the requisite manner.

5. Conclusion

The moral permissibility of actions, I have granted, is based in part on how valuable their consequences are. Having maximally good consequences, however, is, assuming a “standard” theory of value, neither necessary nor sufficient for being permissible. It is not necessary, because only satisfaction is required. It is not sufficient, because some actions with maximally good consequences are impermissible because they do not adequately respect the normative separateness of persons.

Certain consequentialist theories with “non-standard” theories of value can avoid some of the problems that I have raised. First, the problems stemming from the requirement to maximize can be largely overcome, if the theory of value is relatively coarse-grained or incomplete. This avoids the problem of insufficient moral freedom. If value is not agent-disfavoring, then such theories also avoid the problem of excessive required sacrifice. Such theories still fail to recognize adequate supererogation, but this is the least powerful objection to maximizing theories. Second, the problems stemming from the absence of any constraints protecting the normative separateness of persons can also be largely overcome, if the theory of value focuses on rights infringement and gives lexical priority to earlier infringements. The important issue of whether *victim-focused* infringement-minimizing infringements are permissible was briefly addressed, but not resolved. Even if such infringements are permissible, non-minimizing infringements are still impermissible.

In short, the value of consequences is at least sometimes an insistent moral reason for action. Best support by insistent moral reasons, however, is not required for permissibility, and the value of consequences is not the only insistent moral reason. Maximizing consequentialism,

at least in its standard forms, should therefore be rejected.²⁶

PETER VALLENTYNE

References:

- Brennan, S. (1995). "Thresholds for Rights". *Southern Journal of Philosophy*, 33, 143-68.
- Brock, D. (1991). "Defending Options". *Philosophy and Phenomenological Research* 51, 909-913.
- Cummiskey, D. (1996). *Kantian Consequentialism*. Oxford: Oxford University Press.
- Hooker, B. (2000). *Ideal code, real world*. Oxford: Oxford University Press.
- Kagan, S. (1989). *The limits of morality*. New York: Oxford.
- Kagan, S. (1998). *Normative ethics*. Boulder: Westview Press.
- Kamm, F. (1992). "Non-consequentialism, the person as end-in-itself, and the significance of status". *Philosophy and Public Affairs*, 21, 354-389.
- Kamm, F. (1996). *Morality, mortality II*. New York: Oxford University Press.
- Kamm, F. (2000). "Non-Consequentialism". In H. LaFollete (ed.). *The Blackwell guide to ethical theory* (pp. 205-26). Oxford: Blackwell Publishers.
- Lippert-Rasmussen, K. (1996). "Moral status and the impermissibility of minimizing violations". *Philosophy and Public Affairs*, 25, 333-350.
- Lippert-Rasmussen, K. (1999). "In what way are constraints paradoxical?". *Utilitas*, 11, 49-70.
- Mack, E. (1993). "Agent-relativity of value, deontic restraints, and self-ownership". In R.G. Frey and C. Morris (eds.), *Value, welfare, and morality* (pp. 209-32). Cambridge, MA: Cambridge University Press.
- Mack, E. (1998). "Deontic restraints are not agent-relative restrictions". *Social Philosophy and Policy*, 15, 61-83.
- McNaughton, D. and Rawling, P. (1993). "Deontology and agency". *Monist*, 76, 81-100.
- Mulgan, T. (2001). *The Demands of Consequentialism*. Oxford University Press.
- Murphy, L. (2000). *Moral demands in non-ideal theory*. Oxford: Oxford University Press.
- Myers, R. (1999). *Self-governance and cooperation*. Oxford: Oxford University Press.

- Portmore, D.W. (2003). "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation". *Ethics*, 113, 303-32.
- Scheffler, S. (1982). *The rejection of consequentialism*. Oxford: Clarendon Press.
- Scheffler, S. (ed.) (1988). *Consequentialism and its critics*. Oxford: Oxford University Press.
- Scheffler, S. (1992). *Human morality*. Oxford: Oxford University Press.
- Sen, A. (1982). "Rights and Agency". *Philosophy and Public Affairs*, 11, 3-37.
- Slote, M. (1989). *Beyond Optimizing*. Cambridge, MA: Harvard University Press.
- Smart, J.J.C. and Williams B., *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973).
- Vallentyne, P. (1987a). "The Teleological/Deontological Distinction," *Journal of Value Inquiry*, 21, 21-32
- Vallentyne P. (1987b). "Utilitarianism and the Outcomes of Actions", *The Pacific Philosophical Quarterly*, 68, 57-70.
- Vallentyne P. (1988a). "Teleology, Consequentialism, and the Past," *Journal of Value Inquiry*, 22, 89-101.
- Vallentyne P. (1988b). "Gimmicky Representation of Moral Theories," *Metaphilosophy*, 19, 253-63.

Further Reading:

- Applbaum, A. (1998). "Are violations of rights ever right?". *Ethics*, 108, 340-66.
- Bennett, J. (1995). *The act itself*. Oxford: Oxford University Press.
- Brook, R. (1991). "Agency and morality". *Journal of Philosophy*, 88, 190-212.
- Hurley, P. (1997). "Agent-centered restrictions: Clearing the air of paradox". *Ethics*, 108,120-46.
- Kagan, S. (1984). "Does consequentialism demand too much? Recent work on the limits of

- obligation". *Philosophy and Public Affairs*, 13, 239-254.
- Kumar, R. (2000). "Defending the moral moderate: Contractualism and common sense".
Philosophy and Public Affairs, 28, 275-309.
- McCarthy, D. (2000). "Harming and allowing harm". *Ethics*, 110, 749-79.
- McMahon, C. (1991). "The paradox of deontology". *Philosophy and Public Affairs*, 20, 350-77.
- Nagel, T. (1979). "The limits of objectivity". In S. McMurrin (ed.), *Tanner Lectures on Human Values* (pp. 75-139). Cambridge: Cambridge University Press.
- Nagel, T. (1991). *Equality and partiality*. New York: Oxford University Press.
- Otsuka, M. (1997). "Kamm on the Morality of Killing". *Ethics*, 108, 197-207.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Pettit, P. (1988). "The Consequentialist Can Recognize Rights". *Philosophical Quarterly*, 38, 42-55.
- Pettit, P. (ed.) (1993). *Consequentialism*. Brookfield: Dartmouth Publishing Company.
- Ridge, M. (2001). "Saving Scanlon: Contractualism and agent-relativity". *Journal of Political Philosophy*, 9, 472-81.
- Scanlon, T.M. (1998). *What we owe to each other*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Scanlon, T.M. (2000). "Intention and permissibility". *Proceedings of the Aristotelian Society, Supplemental Volume*, 74, 301-317.
- Sen, A. and Williams, B. (eds.) (1982). *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.

¹ Note that I formulate the maximization requirement as "no alternative action has more valuable consequences" rather than as "its consequences are at least as valuable as those of any

alternative”. The former is more general in that it covers cases where some consequences are incomparable in value (neither at least as valuable as the other).

² For criticism of rule consequentialism, see Smart’s essay in Smart and Williams (1973). For a state of the art defense of rule consequentialism, see Hooker (2000).

³ A full discussion of the definition of consequentialism would need to address several other issues. For example, are consequences understood as world prospects (roughly: probability distributions over possible world histories), full world histories (including the past), only world futures, or only the avoidable part of world futures. For further discussion of this and other related issues, see Vallentyne (1987a, 1987b, 1988a, 1988b).

⁴ For discussion of (the slightly unusual notion of) agent-relative moral value, see Sen (1982) and Portmore (2003).

⁵ The most influential work defending act consequentialism on the topics of options and constraints is Kagan (1989). Scheffler (1982) is the most influential defense of options (but he rejects constraints. See also Scheffler (1988, 1992) and Kagan (1998).

⁶ See, for example, ch. 8 of Kamm (1996).

⁷ See, for example, Slote (1989).

⁸ For further discussion, see Murphy (2000), Lippert-Rasmussen (1996, 1999) and Kamm (1996; ch. 8).

⁹ *Agent-relativity* exists when permissibility is sensitive (involves an essential reference) to features of the agent (her beliefs, values, relationships, etc.) other than the specification of which choices are feasible (and certain related matters). It holds that at least sometimes the identification of who the agent is in a given choice situation is relevant for determining permissibility. *Agent-neutrality* denies this. Agent-relativity is compatible with the sensitivity

being agent-favoring (as in ethical egoism), agent-*disfavoring* (as in pure ethical altruism), or *agent-indifferent* (as in a theory that favors the agent's relatives, but gives no special preference to the agent's well-being). For a superb discussion of these notions, see, McNaughton and Rawling (1993).

¹⁰ Note that for highly agent-disfavoring theories of value, such as anti-egoism, the problem of excessive require sacrifice can arise even when value is relatively coarse or incomplete.

¹¹ I here use "option" in its broad sense. Many authors use it more narrowly to mean "morally permitted action that does not (perhaps relative to applicable constraints) maximize the value of consequences". (Options in this sense are sometimes called "prerogatives".) In my sense, but not this more restrictive sense, when two actions both maximize value, they are each options. A maximizing theory with a relatively coarse-grained or incomplete value recognizes many options in my sense but none in the narrow sense. I address the issue of options in the narrower sense below when I discuss the problem of insufficient supererogation.

¹² Roughly speaking, both Kamm (2000) and Brock (1991) defend options on the basis of the need to leave adequate moral freedom (personal sovereignty, autonomy). Kamm further argues that (roughly) avoidance of excessive required sacrifice does not justify options on the grounds that (1) we are not permitted to control someone else's life (e.g., who they kiss) even if we care deeply about what they do, and (2) we are not permitted to mistreat people (e.g., kill, assault, or steal from), simply because we would get an enormous benefit from doing so. I fully agree that morality sometimes requires great sacrifice and that options are not justified solely by the need to avoid excessive sacrifice. This does not show, however, that avoidance of frequent excessive required sacrifice is not relevant to the justifications for options. An adequate theory must both leave agents adequate moral freedom and not require excessive sacrifice.

¹³ Not so implausible is the view that certain values are lexically posterior (relevant only for breaking ties) to others and that only the latter are relevant for determining permissibility. Such a view, however, is not really a maximizing theory.

¹⁴ Portmore (2003) argues that maximizing consequentialist theories can accommodate options and supererogation. His defense of this claim, however, involves invoking two values: moral value and all-things considered value, the latter of which is based on both moral value and prudential value for the agent. He defends a theory that holds that an action is permissible if and only if no other feasible action has consequences that are better *both* in terms of moral value and in terms of all things considered value. Such a theory, however, is equivalent to one based on a relatively incomplete theory of single value: one outcome is better than a second if and only if it is better with respect to both moral and all-things-considered value. Thus, Portmore's approach is equivalent to adopting a certain relatively incomplete theory of value.

¹⁵ Satisficing consequentialism is, of course, highly controversial. One objection is that, if it is unconstrained, it permits the mistreatment of others for one's benefit much more frequently than the corresponding maximizing version. In the next section, however, I argued that consequentialism (whether satisficing or maximizing) needs to be constrained by certain rights. For a more general criticism of satisficing consequentialism, see Mulgan (2001).

¹⁶ Nozick (1974, p. 28), for example, emphasizes that a moral concern need not be a moral goal.

¹⁷ Another way of grounding constraints, which I shall not explore, is to hold that there is a constraint against using force against a person to impose a loss of well-being that that person is not required to impose on herself or to submit to. For insightful discussion of this strategy, see Myers (1999).

¹⁸ Throughout, I assume that interest-protecting rights are *direct* in the sense that a given right

protects the holder's interests on *each* occasion, and not *indirect* in the sense of protecting the holder's interests overall (but not necessarily on each occasion). In this direct sense, an interest-protecting right is *waived* whenever infraction does not thwart the holder's interest, even if she has not consented. (Admittedly, this is a non-standard notion of waiving, since no person does the waiving. It is useful, however, since it preserves the parallel with choice-protecting rights in making a feature of the right-holder's person (interests or will) determinative of whether a boundary-crossing is an infringement.)

¹⁹ Victim-focused (personal rights-based) defenses of constraints are given by Kamm (1996, ch.8; 2000), and Mack (1993, 1998).

²⁰ For insightful discussion of rights with thresholds, see Brennan (1995).

²¹ Things are more complex, Kasper Lippert-Rasmussen pointed out to me, if an action can become a rights violation *after* it has been performed: If Evil fires a gun at Victim1, but this does not violate her rights until the bullet strikes her, then the maximizing theory of the text would judge it permissible, indeed obligatory, to throw Victim 2 in front of the bullet, since that rights violation would occur later than the retroactive rights violation of Victim 1 if the bullet strikes her. For simplicity, I ignore this complication.

²² Sen (1982) explores such a possibility.

²³ Kamm (1992, 1996, chs. 8&9; 2000), and Mack (1993, 1998) both argue against the common (and mistaken) supposition that agent-relativity is the ground of constraints on the promotion of the good. They both argue, as have I, that it is rather a kind of agent-neutral concern for victims.

²⁴ It may also be that the normative separateness of persons also allows killing Smith where (1) this saves at least one other person from being killed, and (2) Smith would be killed by someone else if one did not kill him. In such cases, Smith is arguably not being sacrificed for the benefit

of others.

²⁵ See McNaughton and Rawling (1993) and Kamm (1996, ch.8) for insightful discussions.

²⁶ For valuable comments, I thank Jamie Dreier, Brad Hooker, Paul Hurley, Kasper Lippert-Rasmussen, Piers Rawling, and Bill Shaw.