

CAN AN ACT-CONSEQUENTIALIST THEORY BE AGENT RELATIVE?

Douglas W. Portmore

One thing...all [consequentialist theories] share...is a very simple and seductive idea: namely, that so far as morality is concerned, what people ought to do is to minimize evil and maximize good.... On the face of it, this idea, which lies at the heart of consequentialism, seems hard to resist. For given only the innocent-sounding assumption that good is morally preferable to evil, it seems to embody the principle that we should maximize the desirable and minimize the undesirable, and that principle seems to be one of the main elements of our conception of practical rationality.... Nevertheless, consequentialism has implications that appear to sharply conflict with some of our most firmly held moral convictions. ...[And] non-consequentialists...argue that consequentialism can at best be reconciled not with our actual values, but only with their simulacra. To reflect our actual values, they claim, a moral view must include what have recently come to be called *agent-relative* elements, which are incompatible with consequentialism.

—Samuel Scheffler¹

The above nicely summarizes the current impasse in the debate over consequentialism. On the one hand, consequentialism is very compelling in that it holds that it is always morally permissible to bring about the best available state of affairs—whereas deontology seems almost paradoxical in that it insists that it is wrong to commit a right's violation even in order to minimize such violations overall. But, on the other hand, consequentialism seems to have some very counter-intuitive implications, implications that can only be avoided by a normative theory that incorporates agent-relative elements. However, all this assumes that we cannot have a theory that incorporates both these agent-relative elements and the idea that it is always permissible to bring about the best available state of affairs. In other words, the assumption is that consequentialism—the view that agents ought to minimize evil and maximize the good—goes

hand in hand with agent neutrality. In this paper, I argue that it is a mistake to equate consequentialism with agent-neutrality.

I. INTRODUCTION

A theory is agent neutral if it gives every agent the exact same set of aims and agent relative otherwise.² Thus utilitarianism is agent neutral, for it gives every agent the same single aim: the maximization of aggregate welfare. Kantianism, on the other hand, is agent relative; it gives different agents different aims. For instance, it gives me the aim that I not break *my* promises but you the aim that you not break *your* promises. Of course, we are both bound by the same categorical imperative: “never break one’s promises.” But this imperative generates different aims for different agents, because it contains the indexical “one’s promises,” which refers to my promises when addressed to me but to your promises when addressed to you.

This distinction between agent-relative and agent-neutral theories is now recognized as being one of the most important in normative ethics, for only an agent-relative theory can accommodate the basic features of commonsense morality: *options*, *constraints*, and *special obligations*.³ Consider, for instance, that a theory would have to be agent relative in order to accommodate a constraint against murder, for such a constraint would prohibit the commission of murder even for the sake of preventing numerous others from committing comparable murders. The best an agent-neutral theory can do, on the other hand, is give everyone the aim of minimizing murders; yet this would permit the commission of murder for the sake of minimizing murders overall, which is just contrary to the very nature of a constraint.⁴

Given the importance, then, of agent relativity in accommodating our commonsense moral intuitions, it is important to ask whether an act-consequentialist theory can be agent relative.⁵ If it cannot, there can be no hope of ever reconciling act-consequentialism (hereafter, simply “consequentialism”) with our commonsense moral intuitions.⁶ (I’m assuming, of course, that it

would count in a theory's favor to be in accord with our most basic commonsense intuitions.⁷) Now it seems that the only hope for success along these lines lies with the possibility of combining consequentialism with some sort of agent-relative theory of value.⁸ If such a union is possible, then perhaps consequentialism can accommodate most, if not all, of our commonsense moral intuitions. For consider the theory obtained by applying the consequentialist principle "act always so as to maximize value" to an axiology where both of the following are true: (1) the value of a state of affairs in which a murder has been committed is agent relative such that its value varies according to whether or not the evaluator is the murderer, and (2) the disvalue in an agent committing murder herself is, from her position (that of the agent), greater than the disvalue in numerous others committing comparable murders.⁹ On this theory, it would be wrong to commit murder even in order to minimize murders overall, and it would be wrong on consequentialist grounds: the state of affairs in which the agent commits murder herself is, from her position, worse than the one in which numerous others commit comparable murders.¹⁰

Of course, many philosophers would deny that a consequentialist theory can be agent relative. In this paper, I will examine some of the reasons why philosophers have come to this conclusion and argue that they all prove inadequate. I will then attempt to resolve the issue by determining what it is that is distinctive about consequentialism. It is not, as many suppose, a particular view about what's right and wrong. For, as I will show, consequentialists and non-consequentialists can agree that agents should always act so as to promote value.¹¹ What is distinctive about consequentialism, then, is its acceptance of the principle "act always so as to promote value" as a *fundamental* moral principle—that is, as a principle that does not derive from any other substantive moral principle. Or so I will argue. Lastly, I will show how this distinctive feature of consequentialism can be combined an agent-relative theory of value to yield a promising theoretical alternative to both deontology and utilitarianism.

II. MINIMIZING VIOLATIONS AND THE CONSEQUENTIALISM/NON- CONSEQUENTIALISM DISTINCTION

Let us call a right's violation that prevents numerous others (i.e., two or more) from committing comparable right's violations a *minimizing violation*.¹² In the last section, I suggested that consequentialism could accommodate the intuition that it is impermissible to commit minimizing violations so long as it adopts a suitable, agent-relative theory of value. But to do so requires the consequentialist to include as one of the consequences of committing a minimizing violation that the agent has herself committed a right's violation. Only then can the consequentialist claim that the consequences of committing a minimizing violation are worse than not doing so. Therefore, the consequentialist must adopt a rather broad conception of an act's consequences so as to include the commission of the act itself as one of its consequences. However, some philosophers deny that the act itself is a constitutive part of the state of affairs it produces. But this, I think, is a mistake resulting from philosophers too closely associating consequentialism and utilitarianism in their minds. This association has led many philosophers to mistake features of utilitarianism for features of consequentialism. For instance, prior to Michael Slote's influential work,¹³ it was common to view maximizing rationality as definitive of consequentialism. Yet nowadays most philosophers acknowledge the possibility of *satisficing* consequentialism. Similarly, many philosophers have been led to mistakenly think that consequentialism is only concerned with the value of outcomes, and not with that of the acts that produce them, given that utilitarianism holds that the acts themselves (viewed independently of their outcomes) can be neither good nor bad. But utilitarianism views the acts themselves as evaluatively neutral, not because it is consequentialist, but because it is welfarist. Given welfarism, all that matters intrinsically is the welfare of individuals. As a result, nothing besides welfare matters intrinsically, thereby excluding the evaluative relevance of the acts themselves, as viewed independently of their effects on the welfare of individuals.¹⁴

So it seems a mistake to force upon the consequentialist an arbitrarily narrow construal of an act's consequences. Besides, even some of consequentialism's most ardent critics accept that consequentialism is entitled to a very broad construal of an act's consequences.¹⁵ For instance, consider what Bernard Williams has to say about consequentialism in his famous essay "A Critique of Utilitarianism":

Standardly, the action will be right in virtue of its causal properties, of maximally conducing to good states of affairs. Sometimes, however, the relation of the action to the good state of affairs may not be that of cause to effect—the good state of affairs may be constituted, or partly constituted, by the agent's doing that act....¹⁶

Aside from construing an act's consequences too narrowly, the most common reason for denying the possibility of consequentialism accommodating the impermissibility of minimizing violations is the alleged synonymy of "agent neutrality" and "consequentialism."¹⁷ For instance, David McNaughton and Piers Rawling have claimed that the distinction between consequentialism and non-consequentialism is best drawn in terms of the distinction between agent-neutral and agent-relative theories.¹⁸ On their view, a theory is consequentialist if and only if it is agent neutral and non-consequentialist if and only if it is agent relative. However, as Frances Howard-Snyder points out in her paper "The Heart of Consequentialism,"¹⁹ it is a mistake to draw the distinction between consequentialism and non-consequentialism in this way, for not all agent-neutral theories are consequentialist.²⁰ Imagine, for instance, a moral theory that gives every agent the same solitary aim of ensuring that the streets are paved with gold.²¹ This theory is clearly agent neutral since it gives every agent the same single aim (and thus the same set of aims). But is it consequentialist? McNaughton and Rawling are committed to answering yes, necessarily so, for, as they define consequentialism, all agent-neutral theories are consequentialist. But suppose that this theory holds that the state of affairs in which the people are fed is better than the one in which the streets are paved with gold. And suppose that,

nevertheless, the theory still insists that the right thing to do is to ensure that the streets are paved with gold even where this precludes feeding the people. Surely, in this case, the theory would be non-consequentialist, for, as Howard Snyder rightly points out, a consequentialist theory can never require an agent to produce a worse state of affairs rather than a better.²² Thus an agent-neutral theory needn't be consequentialist.

This alone is enough to show that it is a mistake to equate the consequentialism/non-consequentialism distinction with the agent-neutral/agent-relative distinction. But it is still an open question whether a consequentialist theory can be agent-relative. And the answer to this question is what the possibility of reconciling consequentialism with commonsense morality hinges upon. Howard-Snyder's own view is that a consequentialist theory cannot be agent relative. She believes that at the heart of consequentialism is the idea that "the better a state of affairs *from an agent-neutral point of view*, the more moral reason an agent has to produce it."²³ This, she thinks, is what all consequentialists accept and all non-consequentialists deny.²⁴ Howard-Snyder calls this idea "BETTER."

But is BETTER really what's at the heart of consequentialism? Consider the following theorist (call him Theo). Theo accepts a rather peculiar sort of divine-command theory. He accepts all of the following:

- (1) Morality is God's creation such that there are no moral standards independent of God's will.
- (2) An act is morally permissible if and only if it conforms to God's will.
- (3) God's will is that we all do our best to maximize aggregate pleasure.
- (4) The only thing valuable for its own sake is pleasure and the absence of pain.

Theo is, I will argue, a non-consequentialist who accepts BETTER. I grant that Theo's theory is not very plausible. But it is enough for my purposes that it contains no inconsistencies. For all

that I need to show is that it is logically possible for a non-consequentialist to accept BETTER. This by itself is sufficient to prove that BETTER is not definitive of consequentialism.

Theo accepts BETTER, for he believes that the better a state of affairs (the more aggregate pleasure it contains), the more moral reason an agent has to produce it. He believes, just as all (maximizing) consequentialists do, that agents should always bring about the best available state of affairs. But, unlike the consequentialist, his belief that agents should always bring about the best available state of affairs is derivative of another, more fundamental, moral principle: "act in accordance with God's will"—see (2). This principle in conjunction with his view about what God wills us to do—see (3)—generates the prescription: "act so as to maximize aggregate pleasure." For the consequentialist, on the other hand, the belief that agents should always bring about the best available state of affairs is non-derivative. It is for the consequentialist the most fundamental of all moral principles, the one from which all others derive.

Now the question is, Is Theo a consequentialist? It seems not. For he doesn't believe that there is any necessary connection between what's right and what's valuable. It is, on his view, just a coincidence that God wills us to do what maximizes value. In fact, God could change his mind and will that we pave the streets with gold, in which case, we would be required to ensure that the streets are paved with gold even where this means doing less than we can to promote value. Of course, Theo does, as it happens, accept a theory that is extensionally equivalent to hedonic act-utilitarianism (hereafter, "utilitarianism"); this, however, is not sufficient to make him a utilitarian. For although extensionally equivalent, Theo's theory and utilitarianism are importantly different. According to utilitarianism, it is wrong to do less than one can to promote aggregate pleasure because that entails producing a sub-optimal state of affairs. Theo, on the other hand, thinks that it is wrong to do less than one can to promote aggregate pleasure, not because it involves producing a sub-optimal state of affairs, but because it involves disobeying the will of God. This means that the two theories have different truth conditions. Utilitarianism

is correct only if hedonism is, whereas Theo's theory does not depend on the correctness of hedonism—if God wills it, agents should, on Theo's view, maximize aggregate pleasure whether there are other things of non-instrumental value or not.²⁵ In contrast, Theo's theory depends on God's will whereas utilitarianism does not—after all, an atheist can be a utilitarian.

But besides the two theories having different truth conditions, there is another important difference between the two, one that involves the fact that consequentialists are constrained in a way that non-consequentialists are not. A consequentialist can prohibit only those acts (i.e., act tokens) that result in a sub-optimal state of affairs. Thus the consequentialist can accept, for instance, an absolute constraint against promise-breaking (the act-type) only if she adopts a theory of value according to which breaking a promise always results in a worse state of affairs than keeping it. Such a theory of value would have to include the following four claims: (1) the value of a kept promise is lexically prior to all other values, (2) the disvalue of a broken promise is agent relative such that the extent of its disvalue varies according to whether or not the evaluator is the promise breaker, (3) the agent-relative disvalue in breaking a promise is, from the agent's perspective, greater than the disvalue in any number of other agents breaking their promises, and (4) the agent-relative disvalue of a broken promise is "moment relative"²⁶ such that the disvalue in breaking a promise now is, from the agent's present position, greater than the disvalue in her breaking any number of other promises in the future. Now I am not making any assertion about the plausibility of such an axiology. The point is only that if such a theory of value were correct, then consequentialism would include an absolute constraint against promise-breaking. Note, then, how consequentialism and non-consequentialism differ. The consequentialist can accept such a constraint only if she also accepts the four claims above. The non-consequentialist, on the other hand, is not constrained in this way. She can accept an absolute constraint against promise-breaking without having to accept these four claims. Unlike

the consequentialist, the non-consequentialist can accept that it is sometimes wrong to bring about the best available state of affairs.

So we have seen that a utilitarian and a divine-command theorist can accept extensionally equivalent theories, as both Jeremy Bentham and Theo do. What's more, we can also construct versions of virtue ethics and contractarianism that are extensionally equivalent to utilitarianism. Take, for instance, the virtue ethicist who accepts the fundamental moral principle: "act as a virtuous agent would characteristically do in the circumstances."²⁷ When this principle is coupled with the view that what it is characteristic for virtuous agents to do in all circumstances is to maximize aggregate pleasure, we arrive at a view that is extensionally equivalent to utilitarianism. Similarly, we can generate an extensionally equivalent contractarian position by combining the contractarian principle "act in accordance with those rules that would be chosen by rational agents behind the veil of ignorance" with the view that what rational agents would try to do in such a situation is maximize their expected utilities. The upshot of all this is that we cannot distinguish various normative theories in terms of which acts they take to be right and wrong. Instead, we must distinguish various normative theories in terms of what moral principles they take to be fundamental. And, therefore, what distinguishes the consequentialist from the non-consequentialist is the acceptance of the principle "act always so as to promote value" as a fundamental moral principle. So the consequentialist is one who accepts the principle "always act so as to promote value" *as a fundamental moral principle*. And the non-consequentialist is one who either rejects this principle or accepts it as being derivative of some other, more fundamental, moral principle.

III. GETTING CLEAR ON THE DISTINCTION

I have suggested that we must distinguish consequentialists from non-consequentialists in terms of whether or not they accept the principle "act always so as to promote value" as a fundamental

moral principle, that is, as a principle that does not derive from any other moral principle. Strictly speaking, however, every substantive moral principle derives from the principle "act morally." But the principle "act morally" is a formal principle, and just as all theories of rationality endorse the formal principle "act rationally," all moral theories endorse the formal principle "act morally."²⁸ Moral theories, then, differ according to which substantive moral principles they take to be fundamental. So when I say that all consequentialists take the principle "act always so as to promote value" to be fundamental, or non-derivative, I mean non-derivative of any other *substantive* moral principle.

I should also make clear that my distinction allows for the possibility that the consequentialist takes the principle "act always so as to promote value" to be derivative of certain meta-ethical claims—it just can't be derivative of any other substantive *moral* principle. For instance, someone might come to accept consequentialism on the basis of the following meta-ethical claims: (1) there is a conceptual connection between morality and reasons for action such that an agent is morally required to ϕ if and only if she has a decisive reason to ϕ and (2) there is a conceptual connection between reasons for action and the value of states of affairs such that an agent has a decisive reason to ϕ if and only if ϕ ing would produce the best available state of affairs. These two imply (3) that an agent is morally required to ϕ if and only if ϕ ing would produce the best available state of affairs. And so we see that maximizing consequentialism can be derived from certain meta-ethical claims.²⁹

Now having said this, one might wonder whether Theo is just a consequentialist who differs from other consequentialists in holding a different meta-ethical position. Theo does accept a certain meta-ethical position, namely, that morality is God's creation. But this is not what makes Theo a non-consequentialist. Theo is a non-consequentialist, because he accepts a different normative position than consequentialists do. That is, Theo thinks that one should always act in accordance with God's will. This normative position is not just another way of expressing the

meta-ethical view that morality is God's creation, for one can accept that an act is morally permissible if and only if it conforms with God's will but deny that morality is God's creation. For instance, one might hold the standards of morality to be independent of God's will and accept the principle "act in accordance with God's will" on the grounds that that God is all good and an all-good god would always will that one does what's right and not what's wrong.

IV. AGENT RELATIVITY AND ACT-CONSEQUENTIALISM

For all that I have said so far I still haven't proven that a theory can be both agent-relative and consequentialist. Although I have shown that what is definitive of consequentialism is a fundamental commitment to the principle "act always so as to promote value," there remains the possibility that agent-neutrality is also definitive of consequentialism. In this case, the consequentialism/non-consequentialism distinction would be a bipartite distinction.³⁰ But even if I have not proven that a theory can be both agent relative and consequentialist, I have at least undermined two of the most common reasons for thinking that it cannot: I have argued that it is a mistake both to exclude the act itself from its consequences and to equate that the agent-neutral/agent-relative distinction with the consequentialism/non-consequentialism distinction. Yet, admittedly, there are some philosophers who avoid both mistakes and still insist that agent neutrality is definitive of consequentialism. For instance, both Shelly Kagan and John Broome call theories that assess the rightness of acts solely on basis of their propensity to promote value "teleological" as opposed to "consequentialist."³¹ They recommend that we reserve the term "consequentialism" for those theories that are both teleological and agent neutral. And they both appeal to common usage in support of their terminology. And, admittedly, consequentialism is most often used to refer exclusively to agent-neutral theories—though there are some notable exceptions.³² However, the Kagan/Broome terminology can itself be misleading, because the expression "teleological ethics" is often given a much broader reference. As it was first

introduce, “teleological ethics” was meant to denote only those theories that “hold that the rightness or wrongness of an action is always determined by its tendency to produce certain consequences which are intrinsically good or bad.”³³ But, since then, “consequentialism” has come to supplant the more cumbersome “teleological ethics.” And, when employed nowadays, the expression “teleological ethics” is more often used in a broader sense so as to include not just consequentialism but also virtue ethics. For instance, teleological theories are often defined as those that take goodness to be prior to rightness. But, on this definition, virtue ethics is a form of a teleological ethics because it takes goodness to be prior to rightness. For, according to virtue ethics, an act is right if and only if it is what an agent of good character would do in the circumstances, and so one must first determine what constitutes a *good* character before determining what the right thing to do is.³⁴

Another problem with the Kagan/Broome terminology is that it seems arbitrarily restrictive. After all, we would think it arbitrary to restrict (by definitional means) divine-command theorists to a particular view about God’s will, and, likewise, arbitrary to restrict (by definitional means) contractarians to a particular view about what contractors would agree to. But isn’t it just as arbitrary to restrict consequentialists to a purely agent-neutral theory of value? If no particular substantive view is definitive of consequentialism, then what grounds are there for limiting (by definitional means) the sorts of substantive positions available to the consequentialist? Nevertheless, even if there is no legitimate reason for taking agent neutrality to be definitive of consequentialism, there might be other grounds for restricting the consequentialist to a purely agent-neutral theory of value. To illustrate, consider the following non-arbitrary reason for restricting the divine-command theorist to a particular view about God’s will. If divine command theory is defined as the view that takes as fundamental the principle “act always so as to be in accord with God’s will,” then there may seem to be nothing to prevent a divine-command theorist from adopting the view that God’s will is malicious. Such a position is not ruled out by

the definition of divine-command theory. However, it might be ruled out by the definition of God. If God is, by definition, all-good, then we can rule out all substantive versions of divine-command theory that take God's will to be malicious. But although we might legitimately confine the divine-command theorist to the view that God's will is benevolent on the grounds that God is, by definition, omni-benevolent, the presumption should always be that theorists are free to apply their fundamental principle(s) as they see fit. This means that the burden of proof falls upon those who would restrict the consequentialist to a purely agent-neutral theory of value to provide some sort of rationale for doing so. Let me now attempt to preempt those who might try.

One reason for thinking that consequentialists should be confined to a purely agent-neutral theory of value has to do with a certain conception of intrinsic value. Consequentialism is often defined as the view that takes the right act to be the one that maximizes (or, at least, promotes) *intrinsic* goods. But to say that something is intrinsically good is to say that the source of its goodness lies entirely with its own intrinsic nature and not with its relations or circumstances. As Moore puts it,

To say that a kind of value is 'intrinsic' means merely that the question whether a thing possesses it, and in what degree it possesses it, depends solely on the intrinsic nature of the thing in question.³⁵

Agent-relative value, then, must be a type of extrinsic value, for its value depends upon the relation of the evaluator to the state of affairs being evaluated. So if consequentialism must be defined in terms of intrinsic goods and not extrinsic goods, then agent neutrality is definitive of consequentialism. This, then, seems to be a non-arbitrary reason for restricting the consequentialist to a purely agent-neutral theory of value. However, this rests on a mistake. Consequentialism has come to be defined in terms of intrinsic goods only as a result of a conflation of what are in fact two separate distinctions in goodness.

One is the distinction between things valued for their own sakes and things valued for the sake of something else—between ends and means, or final and instrumental goods. The other is the distinction between things which have their value in themselves and things which derive their value from some other source: intrinsically good things versus extrinsically good things.³⁶

When these two distinctions are kept separate, it becomes clear that consequentialism should be defined in terms of final goods, not intrinsic goods. For if we were to define consequentialism in terms of intrinsic goods, we would exclude from the consequentialist's consideration a whole class of goods, namely, extrinsic goods. Yet the consequentialist is always concerned with whether or not something is good. The source of that goodness—whether it derives from its own intrinsic nature or some other source—is irrelevant. On the other hand, if we define consequentialism in terms of final goods, we do not exclude any class of goods from consideration. Instrumental goods will be accounted for to the full extent of their value, for instrumental goods are valuable only in so far as they are means to some final good. Thus, when defined as the view that takes the right act to be the one that maximizes final goods, consequentialism is able to take full account of all types of goods.

Important, then, is the fact that agent-relative goods can be valued for their own sakes. For instance, it might be that the value of Smith's happiness is agent-relative such that it is greater for a friend than for a stranger, but this is compatible with both the friend and the stranger valuing Smith's happiness for its own sake. So although consequentialism would be rightly restricted to a purely agent-neutral theory of value if consequentialism were only concerned with intrinsic goods, we find that that is not the case.

Another reason for thinking that consequentialism should be confined to a purely agent-neutral theory of value is the thought that consequentialism must be so restricted in order to remain consistent with the objective, categorical, and universal nature of morality. It is common

to view morality as a system of categorical imperatives such that an agent's duties do not depend on her subjective desires.³⁷ On this view, an agent's obligation to keep her promise is not contingent upon her having some subjective end that would be furthered by keeping that promise. Now it is true that if applied to a certain sort of agent-relative theory of value, one where the value of a state of affairs varies depending upon the evaluator's subjective desires, consequentialism would generate different obligations for agents with differing desires. We might, then, arrive at some form of ethical egoism. But we needn't see the relativity of value as being attached to an agent's subjective desires. We can, alternatively, see the relativity of value as being attached to an agent's position. We obtain such a theory by applying the principle "act always so as to maximize value" to a version of *evaluator relativism*³⁸ that relativizes the value of states of affairs to certain objective positions, like that of agent or bystander.³⁹ Suppose, for instance, that the disvalue of the state of affairs in which an agent has committed murder for the sake of preventing five others from committing comparable murders varies according to whether or not the evaluator is the agent in question.⁴⁰ Suppose that from the agent's position the state of affairs in which she commits murder for the sake of preventing the five others from committing murder (call this S_1) is worse than the one where she refrains from committing murder and so allows the five others to each commit murder (call this S_2)—worse, perhaps, because agents bear a special responsibility for their own actions. But now suppose that from the position of an innocent bystander, S_2 is worse than S_1 since there are more murders in S_2 . When this version of evaluator relativism is combined with the principle "act always so as to bring about the best available state of affairs," we obtain categorical imperatives. For instance, those in a position to prevent five others from committing murder by committing murder themselves should refrain from doing so irrespective of their own personal interests or desires.⁴¹ For, on the given axiological view, the disvalue of a murder would not depend on the murder being detrimental to the evaluator's interests or on the evaluator lamenting the victim's death. It would instead

depend solely upon the position of the evaluator: agent or bystander. So the evaluator cannot choose which state of affairs, S_1 or S_2 , is more valuable. For although the correct ranking of the two states of affairs will vary according to the evaluator's position, the ranking is fixed and unvarying for each position. Therefore, S_1 is, objectively speaking, better than S_2 from the position of an innocent bystander. And S_2 is, objectively speaking, better than S_1 from the position of the agent.^{42,43}

So I have shown that what may have initially seemed to be legitimate reasons for restricting the consequentialist to a purely agent-neutral theory of value turn out to be inadequate. But now in absence of any legitimate reason for restricting the consequentialist to a purely agent-neutral theory of value, we should conclude that it would be arbitrary to do so.

V. CONCLUSION

I have argued that there is no good reason for supposing that a normative theory cannot be both consequentialist and agent relative. The obvious response to this is to point out that “consequentialism” is a term of art, and so one can just stipulate that s/he will only use the term to denote a certain subset of agent-neutral theories. And I have no problem with those who make such a conscious stipulation. However, what I have tried to combat in this paper is the tendency among many philosophers to equate consequentialism with agent-neutrality and non-consequentialism with agent-relativity. It is important combat this tendency, because the conflation of the consequentialism/nonconsequentialism and the agent-neutral/agent-relative distinctions has had some rather unfortunate effects on twentieth century moral philosophy. First, it has prevented moral philosophers from exploring a new and interesting theoretical possibility: agent-relative consequentialism. (For the last two centuries, ethicists have focused, almost exclusively, on just two theoretical possibilities: deontology (i.e., agent-relative nonconsequentialism) and utilitarianism (i.e., agent-neutral consequentialism)—leaving the

prospects for a hybrid theory completely unexplored.^{44, 45}) Second, the conflation of these two distinctions has prevented us from breaking the “spell” that utilitarianism has had over most of modern moral philosophy. As a number of philosophers have noted with curiosity, utilitarianism has tended to haunt even those who are convinced that it is wrong. Something about it is very compelling, if not “spellbinding.”⁴⁶ Nevertheless, most insist that utilitarianism is wrong, for too many of its implications conflict with our deeply held moral convictions. Indeed, it is because of its counter-intuitiveness that Bernard Williams predicted in 1973 that “The day cannot be too far off in which we hear no more about it.”⁴⁷ Yet decades later utilitarianism continues to haunt us—non-utilitarians still feel compelled to defend against it. But why has utilitarianism had this spellbinding force? What about it is so compelling? It seems to be the thought that it can never be wrong to bring about the best available state of affairs.^{48, 49} For how could it be wrong to produce what is admittedly the best available alternative?⁵⁰

So, on the one hand, we want to reject any view that has the counter-intuitive implications associated with utilitarianism. But, on the other hand, it seems that we should accept the idea that it is always permissible to pursue the best available state of affairs. Yet, as most see it, one cannot do both. For it may seem that the following are true: (1) only an agent-relative theory can avoid the counter-intuitive implications associated with utilitarianism, (2) only a consequentialist theory can endorse the idea that it is always morally permissible to produce the best available state of affairs, and (3) all consequentialist theories are, by definition, agent neutral. Thus it seems that we are faced with a rather unattractive dilemma: either bite the bullet and accept the counter-intuitive implications associated with utilitarianism or give up the compelling idea that it is always permissible to bring about the best available state of affairs. But, as I have shown, this is in fact a false dilemma. Agent-relative consequentialism shares with all consequentialist theories the idea that it is always permissible to bring about the best available state of affairs—it differs from agent-neutral consequentialist theories only in that it denies that there is always a

state of affairs that can be said to be best, period, irrespective of the evaluator's position in relation to that state of affairs. Yet, unlike agent-neutral consequentialism, agent-relative consequentialism can avoid the counter-intuitive implications associated with utilitarianism. As we have seen, it does so provided that a certain theory of value is correct, e.g., one where the disvalue in committing murder is agent relative. Now ultimately we may decide to reject such a theory of value, but it is worth exploring since the result would be a theory that takes what's best about utilitarianism while leaving behind its counter-intuitive implications.⁵¹

Notes:

¹ From the introduction of *Consequentialism and Its Critics*, edited by Samuel Scheffler, (New York: Oxford University Press: 1988), pp. 1-2.

² This distinction comes from Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984)—see p. 27. Parfit, however, is a bit imprecise in his initial formulation of the distinction. He says that an agent-neutral theory gives all agents common aims and that an agent-relative theory gives different agents different aims. But what about a theory that gives every agent a few common aims while at the same time giving some agents different aims? As I draw the distinction, the theory comes out agent relative. Although it gives all agents some common aims, it does not give every agent the same *set* of aims. Later on, it becomes clear that this is also what Parfit has in mind, as he later refers to agent-neutral theories as those that give to all agents *only* common aims—see p. 129.

³ Both constraints and special obligations limit an agent's freedom to pursue the best overall state of affairs (i.e., the best state of affairs from an impersonal standpoint). Constraints prohibit the commission of certain act-types (e.g., murder) even for the sake of preventing numerous others from committing comparable instances of that act-type. Special obligations are duties specific to an individual given her particular circumstances and history. These include duties arising out of past acts (e.g., the duty to keep one's promises) and also those duties that come with occupying certain roles (e.g., professional duties and familial obligations). In contrast to both constraints and special obligations, options do not prohibit an agent from doing what will produce the best overall state of affairs. Options provide agents with the choice of either safeguarding their own interests or sacrificing those interests for the sake of the overall good. In the case of options, agents are permitted to perform a non-optimific act even where that act does not involve the violation of either a constraint or special obligation. (The above is meant only to be a brief sketch of the

three basic features of commonsense morality. For a more thorough and illuminating account, see Shelly Kagan's *The Limits of Morality* (Oxford: Clarendon Press, 1989) and "The Structure of Normative Ethics," *Philosophical Perspectives* 6 (1992): 223-42.)

⁴ See Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 28-33.

⁵ Clearly rule-consequentialism is agent relative. Like Kantianism, rule-consequentialism includes rules containing indexicals, which generate different aims for different agents. And because of its agent-relativity, it is able to accommodate many of our commonsense moral intuitions. But what I wish to explore in this paper is whether an act-consequentialist theory can be agent relative.

⁶ Although indirect consequentialist theories, such as rule-consequentialism and motive consequentialism, have been somewhat successful in accommodating our commonsense moral intuitions, it is worth considering whether act-consequentialism can be reconciled with our commonsense moral intuitions.

⁷ Although this is widely accepted, there are a few detractors. Most notable in recent years is Peter Unger. See his *Living High and Letting Die* (Oxford: Oxford University Press, 1996), and see Frances Kamm's response: "Grouping and the Imposition of Loss," *Utilitas* 10 (1998): 292-319.

⁸ An agent-relative theory of value is one where the value of a state of affairs can vary from one individual to another, and vary not just in terms whether or not it is valuable but also in the degree to which it is valuable.

⁹ I am purposely being vague in not specifying whether, from the agent's position, there is any number of murders committed by others that would outweigh the agent-relative disvalue in an agent committing murder herself. My own opinion is that there is and that this number, whatever it may be, marks the "threshold" beyond which agents are permitted to commit murder for the sake of minimizing murders overall.

¹⁰ Call the state of affairs where the agent has herself committed murder S_1 , and call the state of affairs where numerous others have committed comparable murders S_2 . The claim that S_1 is, from the agent's position, worse than S_2 is meant to be an objective claim. It may seem, however, that it must be a subjective claim since whether or not S_1 is worse than S_2 depends on the position of the evaluator—that is, it depends on whether she is the murderer or just some innocent bystander. There is, after all, a long tradition of taking objective claims to be those whose truth does not depend on either the position or perspective of the assessor. For many philosophers take the objective view to be the "view from nowhere." As Nagel puts it, "a view or form of thought is more objective than another if it relies less on the specifics of the individual's makeup and position in the world..."—see *The View from Nowhere* (Oxford: Oxford University Press, 1986), p. 5. Nevertheless, there would seem to be such a thing as "positional objectivity." Consider the claim "the sun is setting." The truth of this claim is position-dependent, for if two people, on opposite coasts of the United States, assert, at the same time, that the sun is setting, it can turn out that one is correct while the other incorrect. Yet this is not to say that it is a subjective claim. It is obviously not. For whether the sun is setting at a particular place and time does not depend on anything in the mind. Nor does it

depend on anyone observing that the sun setting. The claim "the sun is setting" is true if and only if the sun is indeed setting at the time and place of its utterance. For a more thorough discussion of positional objectivity, see Amartya Sen, "Positional Objectivity," *Philosophy & Public Affairs* 22 (1993): 126-45.

¹¹ The principle "promote value" is meant to be neutral between maximizing and satisficing versions of consequentialism.

¹² I borrow this term from Kasper Lippert-Rasmussen, "Moral Status and the Impermissibility of Minimizing Violations," *Philosophy & Public Affairs* 25 (1996): 334. The phrase "preventative violation" is also used to mean roughly the same thing—see Christopher McMahon, "The Paradox of Deontology," *Philosophy & Public Affairs* 20 (1991): 350.

¹³ I'm referring to his book *Common-Sense Morality and Consequentialism* (London: Routledge & Kegan Paul, 1985).

¹⁴ This point comes from Amartya Sen, "Well-Being, Agency and Freedom," *Journal of Philosophy* 82 (1985): 181-2.

¹⁵ For some supporters of consequentialism who advocate a broad construal of an act's consequences, see John Broome, *Weighing Goods* (Oxford: Blackwell, 1991), 4; Fred Feldman *Utilitarianism, Hedonism, and Desert* (Cambridge: Cambridge University Press, 1997); and David Sosa "Consequences of Consequentialism," *Mind* 102 (1993): 102-22.

¹⁶ See Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982), pp. 1-2, fn. 2.

¹⁷ See, for example, Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989), p. 8; David McNaughton and Piers Rawling, "Honoring and Promoting Values," *Ethics* 102 (1992): 836; and Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982), pp. 1-3.

¹⁸ See "Honoring and Promoting Values," *Ethics* 102 (1992): 836.

¹⁹ In *Philosophical Studies* 76 (1993): 107-29.

²⁰ As a result of her paper, McNaughton and Rawling have backed away from their initial position. They now believe that agent-neutrality is only a necessary, not a sufficient, condition for a theory being consequentialist. Thus a theory is consequentialist only if it is agent neutral—see "Value and Agent-Relative Reasons," *Utilitas* 7 (1995): 35, fn. 11.

²¹ Howard-Snyder, "The Heart of Consequentialism," *Philosophical Studies* 76 (1993): 109.

²² *Ibid.*

²³ *Ibid.*, p. 112.

²⁴ *Ibid.*, p. 110.

²⁵ Theo does accept hedonism, for, as I imagine him to be, he accepts claim (4). But his theory about what makes an act right does not depend on him accepting hedonism or any other theory of value.

²⁶ I borrow this term from Broome, *Weighing Goods* (Oxford: Blackwell, 1991), p. 10.

- ²⁷ Rosalind Hursthouse, “Normative Virtue Ethics” in Roger Crisp (ed.), *How Should One Live?* (Oxford: Clarendon Press, 1996), pp. 19-36.
- ²⁸ This distinction mirrors Derek Parfit’s distinction between formal and substantive aims—see his *Reasons and Persons* (Oxford: Clarendon Press, 1984), p. 3.
- ²⁹ Although there may not be any sharp distinction between meta-ethics and normative ethics, (1) and (2) are quite clearly meta-ethical. They are about conceptual connections between morality and rationality and rationality and value. And (3) is quite clearly a claim of normative ethics. It is a moral principle.
- ³⁰ The utilitarianism/non-utilitarianism distinction is a bipartite distinction. A theory is utilitarian if and only if it is both consequentialist and welfarist. And a theory is non-utilitarian if and only if it is either non-consequentialist or non-welfarist.
- ³¹ See John Broome, *Weighing Goods* (Oxford: Blackwell, 1991), pp. 3-6 and Shelly Kagan, *Normative Ethics* (Boulder, CO: Westview Press, 1998), pp. 63 & 310.
- ³² See Amartya Sen, “Positional Objectivity,” *Philosophy & Public Affairs* 22 (1993): 142-5; James Dreier, “Structures of Normative Theories,” *The Monist* 76 (1993): 22-40; Frances Kamm, *Morality, Mortality*, Vol. II (Oxford: Oxford University Press, 1997), p. 239; Thomas Hurka, *Perfectionism* (Oxford: Oxford University Press, 1993), p. 60; and Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge, Mass.: Harvard University Press, 1993), pp. 30-1.
- ³³ C. D. Broad, *Five Types of Ethical Theory* (London: Routledge & Kegan Paul, 1930), pp. 206-7.
- ³⁴ See Justin Oakley, “Varieties of Virtue Ethics,” *Ratio (New Series)* IX (1996): p. 129.
- ³⁵ “The Conception of Intrinsic Value” reprinted in Thomas Baldwin (ed.), *Principia Ethica*, Revised Edition (Cambridge: Cambridge University Press, 1993), p. 286. What Moore means by a thing’s intrinsic nature is quite complicated. But for our purposes it is enough to say that it consists in its non-relational properties.
- ³⁶ Christine Korsgaard, “Two Distinctions in Goodness,” *The Philosophical Review* 92 (1983): 170.
- ³⁷ Even Philippa Foot has recanted her earlier claim that morality is nothing but a system of hypothetical imperatives. See her 1994 recantation, which follows the reprint of her “Morality as a System of Hypothetical Imperatives,” in Darwall, Gibbard, and Railton (eds.), *Moral Discourse & Practice* (New York: Oxford University Press, 1997), pp. 313-322.
- ³⁸ Evaluator relativism holds that the value of certain states of affairs varies depending on who the evaluator is.
- ³⁹ These positions are objective for whether one is an agent, patient, or bystander in regards to a particular act does not depend on one’s subjective beliefs, desires, or interests. One is the agent if she is the one who performed the act; one is a patient if she is someone on whom the act was performed; and one is a bystander if neither of the other two applies.
- ⁴⁰ Assume that each murder is comparable not only in terms of the harm it will cause but also in terms of the motives, intentions, and characters of the agents involved.

⁴¹ On the other hand, an innocent bystander who could prevent someone else from committing murder in order to prevent five others from committing murder should refrain from doing so, since S_1 is better than S_2 from the position of bystander.

⁴² See fn. 9.

⁴³ Diane Jeske and Richard Fumerton overlook this objective form of agent-relative consequentialism in their article “Relatives and Relativism,” *Philosophical Studies* 87 (1997): 143-157.

⁴⁴ Utilitarianism is an agent-neutral theory that takes the rightness of actions to be purely a function of the value of their resultant states of affairs. Deontology is its converse: an agent-relative theory that holds some acts to be right (or wrong) independent of their propensity to promote value. This preoccupation with just deontology and utilitarianism has left the prospects of a promising hybrid (viz., agent-relative consequentialism) completely unexplored.

On the other hand, the other possible hybrid, agent-neutral nonconsequentialism, seems rather unpromising. For instance, imagine a theory that directs everyone to single-mindedly pursue the state of affairs where all the streets are paved with gold. This theory is clearly agent neutral, since it gives every agent the same single aim. But it might be nonconsequentialist as well. For suppose that the theory holds that the state of affairs where the people are fed is better than the one where the streets are paved with gold and yet insists that the right thing to do is to ensure that the streets are paved with gold even where this precludes feeding the people. Clearly, in this case, the theory would be nonconsequentialist. (This example is borrowed from Frances Howard-Snyder, “The Heart of Consequentialism,” *Philosophical Studies* 76 (1993): 107-29.)

⁴⁵ It is encouraging to see that this is beginning to change as is evident from the recent revival of virtue ethics in the last half of this century. Nevertheless, when it comes to those ethical theories that take acts as opposed to agents as their “primary evaluative focal point,” philosophical discussion still seems to be limited to just these two theoretical possibilities. The expression “primary evaluative focal point” is borrowed from Shelly Kagan’s “The Structure of Normative Ethics,” *Philosophical Perspectives* 6 (1992): 223-42.

⁴⁶ This adjective comes from Philippa Foot, “Utilitarianism and the Virtues,” reprinted in Samuel Scheffler (ed.), *Consequentialism and Its Critics* (Oxford: Oxford University Press, 1988), p. 227.

⁴⁷ See J. J. C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge University Press, 1973), 150.

⁴⁸ Both Philippa Foot and Samuel Scheffler share a similar view about what is so compelling about consequentialism. For Scheffler, it is the idea that “so far as morality is concerned, what people ought to do is to minimize evil and maximize the good”—see his introduction to *Consequentialism and Its Critics* (Oxford: Oxford University Press, 1988): 1. And, for Foot, it is the idea that “it can never be right to prefer a worse state of affairs to a better”—see her “Utilitarianism and the Virtues” reprinted in *Consequentialism and Its Critics*, p. 227. But, on both their views, satisficing consequentialism would lack the compelling

feature that other consequentialist theories share. I prefer, however, to diagnose the compelling feature as something that all consequentialist theories share—i.e., the idea that it can never be wrong to pursue the best available state of affairs.

⁴⁹ Now Philippa Foot has tried to break the “spell” that consequentialism has had over us by arguing that this compelling thought is also what is most radically wrong with consequentialism. She believes that the consequentialist must use the expression “the best state of affairs” in a non-speaker-relative way (that is, in an agent-neutral way) by tacking on “from a moral point of view.” But if we are talking about the best state of affairs from a moral point of view, then she sees no reason for the non-consequentialist to accept such claims as the best state of affairs is the one where an agent commits murder in order to minimize murders overall. For non-consequentialists can claim that comparisons of the value of states of affairs as made from a moral point of view fall solely within the scope of the virtue of benevolence. And since talk of benevolence is limited to acts that are just, it makes no sense to talk about a state of affairs in which a murder has been committed as being the best. Rather, it only makes sense to compare the value of states of affairs in which no unjust act has been committed. As Foot puts it, “sometimes justice will forbid a certain action... and then it will not be possible to ask whether ‘the state of affairs’ containing the action and its result will be better or worse than one in which the action is not done. The action is one that *cannot* be done, because justice forbids it, and nothing that has this moral character comes within the scope of the kind of comparison of total outcomes that benevolence sometimes require”—see “Utilitarianism and the Virtues” reprinted in Samuel Scheffler (ed.), *Consequentialism and Its Critics* (Oxford: Oxford University Press, 1988): 237. Whether we accept the rest of her argument, we might question her initial assumption. Why must the consequentialist use the expression “the best state of affairs” in a non-speaker-relative (i.e., agent-neutral) way? Foot never explains why except to say that moral judgments cannot have a speaker-relative status—see “Utilitarianism and the Virtues” reprinted in Samuel Scheffler (ed.), *Consequentialism and Its Critics* (Oxford: Oxford University Press, 1988): 231. Presumably, the thought is that morality cannot be subjective and that it would be if moral judgments had speaker-relative status. But, as we saw earlier, there is nothing subjective about agent-relative consequentialism despite the speaker-relative status of some of its judgments.

⁵⁰ To see just how compelling the idea that it can never be wrong to produce the best state of affairs really is, consider the deontological analogue of prudence. On such a view, certain acts would be imprudent even though they would bring about what is, from the agent’s position, the prudentially best state of affairs. On this view, certain acts are intrinsically imprudent, imprudent regardless of how prudentially good their consequences. Such a theory seems very odd to say the least. How could it be imprudent to bring about what’s prudentially best? But all non-consequentialist theories share this same oddness. They hold that certain acts are immoral even though they bring about the morally best outcome.

⁵¹ I have benefited from Frances Howard-Snyder’s helpful comments and criticisms on an earlier draft of this paper.

