
UTILITARIANISM AND THE OUTCOMES OF ACTIONS

BY

PETER VALLENTYNE

I. Introduction

Outcome teleological theories judge an action permissible just in case its (objectively determined) outcome (or that of an associated joint action, or rule) is maximally good. Not all teleological theories are outcome teleological: many base the permissibility of actions on the goodness of their (intersubjectively) reasonably anticipatable outcomes, or their (subjectively) anticipated outcomes. Such theories base the permissibility of actions on what the agent believes, or should reasonably believe, the actions' outcomes to be—not on the outcomes themselves.¹

In this paper I shall criticize the conceptions of outcome employed by Marcus Singer and by Donald Regan in recent discussions of outcome teleological theories (and utilitarianism in particular). My aim is not to defend such theories, but rather to illustrate the importance of correctly explicating the notion of outcome both for moral theory and for the theory of rational choice.

II. Outcomes of Actions

The outcome of an action is a state of affairs that in an objective sense would be realized if the action were performed. There are, however, different notions of outcome, based on different further conditions that may be imposed.

In any given choice situation the past (P) is unavoidable. Nothing the agent does can change the past. Furthermore, some parts of the future may be unavoidable (U), in that they will be realized no matter which

feasible action the agent performs. (For example, that a large meteor strikes the ground in China at $t+1$ may be unavoidable relative to the choice situation at t of an agent in the U.S.A.) Other parts of the future may be avoidable (i.e., will not be realized if certain actions are performed). The avoidable future of an action a_1 ($AF(a_1)$) is the most complete entirely avoidable state of affairs that would be realized if a_1 were performed.²

Thus, we have at least the three following notions of outcomes: (1) The *world scenario* of an action is the most complete state of affairs (with no restriction on the times to which it may pertain) that would be the case, if the action were performed. (The world scenario of a_1 is $P\&U\&AF(a_1)$.) (2) The *future* of an action is the most complete state of affairs *not pertaining to the past* that would be the case, if the action were performed. (The future of a_1 is $U\&AF(a_1)$.) (3) The *consequence* of an action is the most complete *entirely avoidable* state of affairs that would be realized, if the action were performed. (The consequence of a_1 is $AF(a_1)$.) Unlike the previous two types of outcomes, the consequence of an action does not include unavoidable states of affairs (i.e., that will be realized no matter what the agent does).³

Although for some purposes it is crucial to distinguish these different types of outcome relations, the criticisms I shall make are not directed at a failure to do so. As I shall explain below, my basic criticism rather concerns a failure to distinguish the notion of outcome from the notion of what happens. First, however, a few remarks are in order concerning some of the notions used in the definitions of the various outcome relations.

The outcome of an action is the most *complete state of affairs* (of a certain sort) that *would* be realized if the action were performed. Three notions deserve comment: (1) The completeness of a state of affairs is to be understood as relative to some implicitly specified conceptual scheme. (2) The "would" ["might"] conditional is to be understood as follows: 'if A were realized, then B would [might] be realized' is true in world w at time t just in case B is true in all [at least one of] the empirically possible world histories of w at t in which A is true. (A world history is empirically possible at t in a world just in case w is the same up to t as the actual history of w , and is compatible with the empirical laws of w .) Thus, the truth of such a conditional does not depend on how things are in other worlds with different pasts or different empirical laws. (3) Both probabilistic (e.g., that there is a 30 per cent chance that it will rain) and possibilistic states of affairs (e.g., that there is a chance that it will rain) are counted as states of affairs. Thus, if it is false that p *would* be realized if ac were performed, but true that p *might* be realized if ac were performed (it depending on other factors), then the outcome of ac does *not* include p , but would include the state of affairs that p is empirically possible. (This, of course, is the usual case, if determinism is false.) If

there are such things as objective probabilities, and I believe that there are, the outcome also includes a state of affairs that ascribes an objective empirical probability to p .

Thus, if there are objective probabilities, the outcome of an action is in effect a probability distribution that ascribes to each non-modal state of affairs its objective probability. If there are no objective probabilities, but only objective possibilities (as governed by current and past state of the world and the laws of nature), the outcome of an action is in effect a "possibility distribution", i.e., a function that ascribes to each non-modal state of affairs its status concerning empirical possibility.⁴

We can now see that the various outcome relations should not be confused with the notion of that which *happens* after an action is performed. Not everything that happens after an action is performed is part of the action's outcome. For example, suppose that if I toss a fair die, there is an objective (ontological, and not merely epistemic) probability of $1/6$ of it coming up on each of 1, 2, . . . 6, respectively, and that I roll the die and it comes up 6. That a 6 comes up is something that happens after I roll the die, but it is not part of the outcome (e.g., future) of my rolling the die. The outcome of my rolling the die would include (entail) the states of affairs that consist of 1, 2, . . . 6 each having an objective probability of $1/6$, but does *not* include the state of affairs that consists of 6 coming up (i.e., having objective probability of 1.0). If (as the example assumes) indeterminism holds with respect to how the die will land if I toss it, then at the time of action there is no fact of the matter as to what will happen (it is not ontologically determined), although there are facts of the matter concerning the objective probabilities of various things happening.

The outcome of an action includes only those states of affairs that *would* be realized if the action were performed. States of affairs that might not be realized if the action were performed, but which (due to the workings of objective chance) are realized, are not part of the outcome of the action; they are merely something that happens.

Many will be puzzled by this claim. Surely, they will say, that a 6 is rolled is part of the outcome of tossing the die. The tossing of the die (let us suppose) increased the chance that a 6 would be rolled, and a 6 was rolled. My reply is that tossing the die guaranteed that there was an objective chance of $1/6$ that a 6 would be rolled, and so that (that the objective chance is $1/6$) is part of its outcome. Tossing the die did not guarantee that a 6 would be rolled, so that (that a 6 is rolled) is not part of its outcome.

To see that my notion of outcome is appropriate for outcome teleological theories, consider a choice situation in which evil aliens have set up a conditional time bomb as follows: If I do not push a certain button, then with certainty Toronto (and nothing else) will be blown up. If I push the button, a fair die is rolled. If the die comes up 1, then with

certainty Toronto will be blown up, and one house in Montreal will be blown up. If the die comes up 2-6, however, then with certainty nothing will be blown up.

Suppose I push the button, it comes up 1, and Toronto and the house in Montreal are blown up. Does that show that the outcome of pushing the button is worse than that of not pushing? No! Clearly, what has *happened*—namely both Toronto and a house in Montreal being blown up—is worse than what would have happened, had I not pushed (just Toronto would have been blown up), but that was just bad luck. Things had a very good ($5/6$) chance of being much better (nothing being blown up). The *outcome* of pushing ($1/6$ chance of Toronto and the house being blown up, $5/6$ chance of nothing being blown up) is (assuming our theory of the good is not extremely risk adverse) better than that of not pushing (the certainty of Toronto and nothing else being blown up). Plausible outcome teleological theories (and standard decision theory) will judge it wrong to not push the button. Even if I push, and both Toronto and the house are blown up, they will still say I did the right thing. Given indeterminism, there is no guarantee that things will turn out as best as possible after an action with a best outcome is performed. Plausible outcome teleological theories judge actions by their outcomes (what they guarantee)—not by everything that happens after they are performed.

Note that if determinism holds, then my account agrees with the outcome-as-what-happens account. For if determinism is true, then at the time of action the most complete state of affairs that *would be realized if* the action were performed just is the most complete state of affairs that *happens after* the action is performed. The two accounts differ only when determinism is false. And as argued above, my account gives a better account in such cases. Indeed it seems highly plausible that the currency of equating outcome with what happens is due to the fact that authors almost always have assumed that determinism is true when discussing the notion of outcome. My account of outcome can thus be seen as an extension of the more common account so as to adequately cover cases where determinism is false.

Let us now see how the new account of outcome can help us assess certain arguments concerning outcome teleological theories.

III. Marcus Singer Against "Actual Consequence" Utilitarianism

Marcus Singer⁵ tells us that Henry Sidgwick and G.E. Moore introduced two important changes in utilitarian (and more generally: teleological) thought. First, Sidgwick (and following him, Moore) explicitly introduced the maximizing formulation of utilitarianism, according to which an action is permissible just in case it has maximally good consequences. Prior to

that time (e.g., on Bentham's and Mill's accounts), the standard form of utilitarianism was non-comparative in that it judged an action permissible just in case it had *good* consequences—even if it did not have maximally good consequences. Second, Moore explicitly introduced the *actual consequence* version of utilitarianism, according to which an action is permissible just in case its *actual* consequences are maximally good. Prior to Moore the usual formulations, according to Singer, were in terms of anticipated or reasonably anticipatable consequences.

With this historical account I have no quarrel. Singer goes on, however, to claim that actual consequence utilitarianism is incoherent. He claims that:

Insofar as it involves comparison of the actual consequences of the act actually performed with the actual consequences of those not performed, it involves something impossible. For no alternative can "turn out best" [T]here can be no actual consequences of actions that are not performed. . . .⁶

Singer claims that actions that are not performed do not have actual consequences, and so any theory that requires the comparison of the goodness of their "actual consequences" is incoherent. His claim applies not only to the consequences of actions in the technical sense introduced above, but also to the other types of outcomes (e.g., world scenarios and futures). For uniformity of terminology with Singer, in this section I shall often focus on the technical notion of consequence, but everything that follows applies equally well to the other notions of outcome.

There seem to be several distinct strands of argumentation that Singer offers in defense of his incoherence claim. One strand rightly claims that we are never in possession of more than a small fraction of the information required to determine whether a given action is permissible according to "actual consequence" utilitarianism, and concludes therefore that the theory is incoherent.⁷ We can certainly grant that "actual consequence" utilitarianism is very difficult (if not next to impossible) to apply correctly, but it is far from clear that the theory is useless as a practical guide to action, let alone incoherent. All it shows is that in applying the theory we will make lots of mistakes. It also suggests that the theory would direct us to generally follow rules of thumb as a practical guide to action. In any case, I do not want to defend here the status of "actual consequence" utilitarianism as a practical guide to action. What concerns me is its status as a theory of what makes actions right and wrong.

Singer argues that "actual consequence" utilitarianism is incoherent as a theory of what makes actions right and wrong on the grounds that actions that are not performed do not have actual consequences. There are two different strands of supporting argument. The less interesting one is that, although actions that are not performed have consequences, they do not

have *actual* consequences, since they are not realized.⁸ There is, of course, a use of 'actual' according to which this is correct, but it trivializes the claim. The role of 'actual' in the expression 'actual consequence' as used in utilitarian theory is to emphasize that what is being referred to is the consequence itself—not what someone believes it to be, or what one might reasonably believe it to be. The adjective 'objective' would be more appropriate than 'actual' for this purpose. "Actual consequence" utilitarianism bases the permissibility of actions on their objective consequences, whether or not those consequences are realized (by the performance of the action). So this line of argument misconstrues the relevant notion of actual consequence.⁹

The more interesting, but nonetheless mistaken, argument in support of the claim that actions that are not performed do not have actual consequences is based on the claim that, assuming indeterminism to be true, at the time of action there is in general no objective fact as to what will happen after any given action is performed.¹⁰ It is only after the action is performed, it is claimed, that there is a fact of the matter. Thus, "actual consequence" utilitarianism, it is claimed, is necessarily backward looking. As a prospective theory of what makes actions right, it is, Singer claims, incoherent.¹¹

Singer is absolutely right that prospective teleological theories that base the permissibility of actions on the goodness of what happens after they are performed are incoherent.¹² The important point to note, however, is that this does not show that "actual outcome", i.e., objective outcome, teleological theories are incoherent. Such theories base the permissibility of an action on the goodness of the most complete state of affairs (of a certain sort) that (objectively) would be realized if the action were performed. Although at the time of action there is in general (assuming indeterminism) no fact of the matter as to what will *happen* after a given action is performed, there is a fact of the matter as to what the most complete possibilistic or probabilistic state of affairs is that would be realized if the action were performed. This is true not only of actions that are actually performed, but also of unperformed but feasible actions. Singer's mistaken conclusion that "actual outcome" utilitarianism is incoherent is based on a mistaken equation of outcome with what happens.

Thus, a useful (although not the intended) reading of Singer's argument is as a *reductio ad absurdum* of equating outcome with what happens. Prospective outcome teleological theories cannot be based on the goodness of what happens after actions are performed. This gives indirect support to my account of outcome.

IV. *Donald Regan on Act Utilitarianism*

In his important book *Utilitarianism and Co-operation*¹³ Donald Regan identifies certain properties that he claims any adequate utilitarian (and more generally: teleological) theory must have, argues that no traditional utilitarian theory (e.g., act, rule, or generalization utilitarianism) has all of these properties, and then defends a new form of utilitarianism, called 'co-operative utilitarianism', on the grounds that it does have all of these properties.

Here I shall consider two claims that he makes about act utilitarianism, and criticize the notion of outcome that he uses in the supporting arguments. The main point of controversy will concern the role of probabilities in the specification of the outcome of an action.

One of the main conclusions of Regan's book is that act utilitarianism does not have a property that he calls 'PropCOP'. A theory has PropCOP just in case its universal satisfaction (by all agents all the time) guarantees the best outcome that any pattern of behaviour by the universe of agents could possibly produce, that is, just in case being a joint action (that is, an n-tuple of individual actions, each having the same time of performance, one for each agent¹⁴) each of the individual action components of which satisfies the theory is sufficient for being a joint action with a best outcome (an outcome that is at least as good as that of any other feasible joint action). Although Regan's claim that act utilitarianism does not have PropCOP is true, the main argument he gives for this claim is flawed. Because the mistake in the argument is an important one, it will be instructive to expose it, even though his conclusion will still stand. We shall see, however, that a second major claim of Regan's can be shown to be false, once the mistaken reasoning is identified.

Let us note first that my account of the notion of outcome applies equally well to joint actions (n-tuples of contemporaneous individual actions, one for each agent). The outcome of a joint action is the most complete state of affairs (of a certain sort) that would be realized if the joint action were performed. As in the case of individual actions, in cases of indeterminism the outcome includes possibilistic and/or probabilistic states of affairs.

Let us start by showing that it is true that act utilitarianism does not have PropCOP. Suppose that there are exactly two agents in the world, Whiff and Poof, and that in a certain choice situation they each have the choice of either pushing a button or not pushing, for each there being an objective probability of 0.3 that he will push, with their actions being causally independent of each other, and with their joint action completely determining what happens (there are no chancy events other than their own behaviour). Suppose further that the goodness of what happens is represented by the following matrix:

		Poof	
		Push	Not-Push
Whiff	Push	10	0
	Not-Push	0	6

In this choice situation act utilitarianism directs each to not push (since $\{[0.3 \times 0] + [0.7 \times 6]\} > \{[0.3 \times 10] + [0.7 \times 0]\}$). Thus, in this situation the universal satisfaction of act utilitarianism (each not pushing) has an outcome (of value 6) that is, not as good as some other outcome that the agents could have jointly achieved (both pushing would produce an outcome of value 10). So act utilitarianism does not have PropCOP.

Regan in effect gives an argument of this sort in support of the claim that act utilitarianism does not have propCOP¹⁵, but he does not treat this argument as his main argument. This is because, although he explicitly recognizes that the outcomes of actions may need to be probabilistic states of affairs, he is not sure that it is appropriate to ascribe objective probabilities to the behaviour of agents.¹⁶

Regan's main argument for the claim that act utilitarianism does not have PropCOP does not involve ascribing objective probabilities to the behavior of agents. With reference to the choice situation described above, except that no assumption is made about the objective probabilities governing their behaviour, he claims that if both do not push, both will satisfy act utilitarianism.¹⁷ Consequently, Regan concludes that act utilitarianism does not have PropCOP, since if both push, the outcome (with value 10) will be better than both satisfying act utilitarianism by not pushing.

The controversial claim here is that both satisfy act utilitarianism if both do not push. His argument in support of this claim is that if Poof does not push, then Whiff's not pushing has a better outcome (6) than his not pushing (0), so Whiff satisfies act utilitarianism by not pushing. Likewise, if Whiff does not push, then Poof's not pushing has a better outcome (6) than his not pushing (0), so Poof satisfies act utilitarianism by not pushing.

What this argument shows is that: (1) *if* the choice situation is such that it is empirically necessary that Poof not push (i.e., it is empirically impossible that he push), then Whiff's not pushing satisfies act utilitarianism; and (2) *if* the choice situation is such that it is empirically necessary that Whiff not push, then Poof's not pushing satisfies act utilitarianism. But any choice situation which satisfies the antecedents of *both* these conditionals, will not provide a counterexample to the claim that act utilitarianism has PropCOP. For in such a choice situation there is only one empirically possible pattern of behaviour (joint action) for the agents, namely both not pushing. Since this is the only empirically possible pattern of behaviour, it trivially has the best empirically possible outcome.

In order for the above choice situation to be a counterexample to the claim that act utilitarianism has PropCOP, it must be empirically possible for each of Whiff and Poof to push and empirically possible for each to not push. Whether such situations are counterexamples depends on the probabilities associated with each action. The above situation will be a counterexample, if for each the probability of pushing is less than $3/8$.

As mentioned above, Regan is unsure about the appropriateness of ascribing objective probabilities to agents, and so wants to formulate his arguments in a way that does not presuppose the appropriateness of such ascriptions. He does this by interpreting act utilitarianism as judging an action permissible just in case, in the circumstances *and given the eventual behaviour of the other agents*, it has a maximally good outcome. Although this is a coherent version of utilitarianism,¹⁸ it is not the form that act utilitarianism usually takes.¹⁹ The usual form is prospective in that it determines the permissibility of actions (based on the goodness of their outcomes) at the time of action. Regan's formulation, however, is not prospective, but only retrospective. According to this formulation at the time of action, there is no fact of the matter as to which actions are permissible. This is because (assuming that the behaviour of the other agents is not strictly determined) at the time of action there is no fact of the matter as to what the eventual behavior of the other agents will be. Only after the other agents have performed their actions is there a fact as to which of the agent's actions had the best outcome on the assumption that the others would behave the way they did.

So, although Regan's formulation of act utilitarianism is coherent, it does not correspond with the usual formulation. The most natural formulation is both realistic (taking the choice situation as it is, not on counterfactual assumptions) and prospective (giving an answer at the time of action to the questions as to which actions are permissible). The most natural version of act utilitarianism requires the agent to respond to the *actual* dispositions of others to behave. If there are no objective probabilities governing this behaviour, such a theory would require the agent to respond to the objective *possibilities* governing the behaviour (e.g., that is empirically possible that X push, but not empirically possible that X swim the Atlantic in one second). Because Regan's version of utilitarianism is neither fully realistic (it ignores the actual dispositions to behave of the other agents) nor prospective (it does not give an answer at the time of action as to which actions are permissible), it is not an intuitively natural version of act utilitarianism.

So far I have endorsed Regan's claim that act utilitarianism does not have PropCOP, but I have criticized his supporting argument. Regan goes on to claim that, although being a joint action, all the individual actions of which satisfy act utilitarianism is not a *sufficient* condition for being a joint action with a best outcome, it is a *necessary* condition. This latter claim is

false of the intuitively most natural version of act utilitarianism. Let me first give a counterexample, and then explain the mistake in Regan's reasoning.

Consider the example given above in which the payoffs are 10 if both Whiff and Poof push, 6 if neither pushes, and 0 otherwise, and in which for each the probability of pushing is 0.3. In this choice situation the joint action with the best outcome has both pushing (with a payoff of 10), and yet, given these probabilities, act utilitarianism directs each not to push (since $|[0.3 \times 0] + [0.7 \times 6]| > |[0.3 \times 10] + [0.7 \times 0]|$). So, being a joint action, each individual action of which satisfies act utilitarianism (in this case neither pushing) is not a necessary condition for being a joint action with a best outcome (in this case both pushing).²⁰

Why does Regan believe otherwise? He reasons as follows.²¹ Suppose, for a reductio, that there is a choice situation in which a joint action with a best outcome (i.e., an outcome that is at least as good as that of any other feasible joint action) has a component action, *ac*, which does not satisfy act utilitarianism. If *ac* does not satisfy act utilitarianism, then there is some alternative feasible action, *ac'*, for its agent to perform that has a better outcome. (So far so good.) But (and here comes the mistake) if that is so, then the joint action cannot have a best outcome. For (Regan mistakenly claims) if *ac'* has a better outcome than *ac*, then the joint action identical to the original except that *ac'* is performed instead of *ac* must have a better outcome than the original joint action. Because this contradicts our supposition that the original joint action has a best (joint action) outcome but has a component action that did not satisfy act utilitarianism, the supposition (he concludes) must be false.²²

The claim made in the second to last sentence is false. From the fact that *ac'* has a better outcome than *ac* in a given choice situation, *the specification of which includes the dispositions to behave (as represented by an objectively determined probability, or possibility, function)* it does not follow that any joint action having *ac'* as a component will have a better outcome than a joint action differing from the original only in that *ac* rather than *ac'* is performed. This is because the outcome of a joint action of which *ac'* (or *ac*) is a component is identical to the outcome of *ac'* (or *ac*) *on the assumption that all the other agents will, with probability one, perform their respective individual actions that are part of the given joint action.* And from the fact that *ac'* has a better outcome than *ac* given the actual dispositions to behave of the other agents, it does not follow that it has a better outcome given the hypothetical assumption that all the other agents will perform the specified actions. This is shown clearly in the example given above. On the assumption that Whiff will perform the action required of him by the best joint action (both pushing), the best action Poof can perform is to push. But given Whiff's *actual* disposition to behave (to push with probability 0.3), the best action Poof can perform is to not push.

In a note²³ Regan anticipates this objection, but his reply is inadequate. He tentatively suggests that the only coherent version of utilitarianism that requires agents to respond to the objective probabilities governing the behaviour of others is one that views agents as having only one feasible action: a mixed action that is the appropriate probability distribution over pure actions. (Pure actions are not viewed as real options, but only as ways that real options can be realized.) On this view, Regan's claim that being a joint action each of the components of which satisfies act utilitarianism is a necessary condition for being a joint action with a best outcome is true. For on this view, there is only one feasible joint action: it consists of each of the uniquely feasible mixed actions of each agent. Thus, trivially a best joint action (there is only one) has components all of which satisfy act utilitarianism (since for each agent the only feasible—and therefore the best—action is a component of the joint action).

The problem with this reply is that there is no reason why requiring agents to respond to the objective probabilities governing the behaviour of others forces one to treat agents as having only the mixed action with the appropriate probability distribution open to them. Just as the "disposition" of a die to land in various ways can be probabilistic without it being that the only one real "option" or "choice" for the die is a mixed "option" consisting of the appropriate probability distribution, so too the dispositions of agents can be probabilistic without their choices so being.

Regan also suggests that we need to reject the claim that act utilitarianism requires an agent to respond to the objective probabilities governing the behaviour of other agents. Of course, we might formulate other versions of utilitarianism that do not have this property, but, if there are such objective probabilities, then, as argued above, the most natural formulation does have this property, and it is with that theory that I am concerned. In any case, even if agents are only required to respond to the objective *possibilities* governing the behaviour of others (e.g., if there are no objective probabilities), it can be shown that natural formulations of utilitarianism are not such that being a joint action each component of which satisfies the theory is necessary for being a joint action that has a best outcome. For example, if (in the absence of objective probabilities) act utilitarianism judged one possibilistic state of affairs better than another just in case the worst non-possibilistic way in which the first might be realized is better than that of the second, the following choice situation shows that it does not have the ascribed property.

		Poof	
		Push	Not-Push
Whiff	Push	10	5
	Not-Push	0	6

In this choice situation the best joint action is that of both pushing, but Poof's pushing does not satisfy act utilitarianism. This is because the goodness value of the outcome of Poof's pushing (the minimum of 10 and 0, i.e., 0) is less than the goodness value of his not pushing (the minimum of 5 and 6, i.e., 5). Thus, on this version of act utilitarianism the best joint action need not be such that its components all have best outcomes.

Of course, the above example depends crucially on how the goodness of possibilistic states of affairs is assessed. A version of act utilitarianism based on a theory of the good that assesses the goodness of possibilistic states of affairs as the *best* way that it could be realized does have the property in question. The point here is that whether act utilitarianism has the property depends on how its theory of the good deals with possibility.

So, even if act utilitarianism does not require agents to respond to the probabilities governing the behaviour of other agents, the most natural version would require them to respond to the possibilities governing the behaviour of others, and, depending on the theory of the good, such a theory need not have the property Regan claims act utilitarianism to have.

V. Conclusion

The outcome of an action is the most complete state of affairs (of various sorts) that (objectively) would be realized if the action were performed. As such it reflects objective probabilities (if there are any) and objective possibilities. The notion of the outcome of an action should not be confused (as in Singer's argument) with the notion of that which happens after the action is performed. As Singer correctly argues theories that base the permissibility of actions on what happens after they are performed are incoherent. Nor should the notion of outcome of an action be confused (as in Regan's arguments) with the notion of the state of affairs that would be realized, given the eventual behaviour of the other agents, if the action were performed.²⁴ Theories that base the permissibility on the latter notion, although coherent, are neither prospective nor realistic.

University of Western Ontario
London, Ontario

NOTES

¹ For further discussion of the differences between outcome teleological theories and other types of teleological theories, see my "The Teleological/Deontological Distinction", *Journal of Value Inquiry*, 21 (1987): 21-32.

² For simplicity I ignore the present, and assume that there are only the past and future. Also, giving an adequate explication of the notion of a state of affairs being *entirely* avoidable is not a simple task. Without resorting to the notion of basic states of affairs, the best explication is that given by J. Howard Sobel, "Utilitarianism: Simple and General", *Inquiry* 13 (1970):394-449, esp. pp.401-403. He does not explicitly offer an explication of the notion of entire avoidability, but it is implicit in his explication of the notion of consequence. For a criticism, see Allan Gibbard, "Doing No More Harm Than Good", *Philosophical Studies* 24 (1973):158-173.

³ Surprisingly few authors have been sufficiently familiar with the different notions of outcome. Notable exceptions include: C.D. Broad, "The Doctrine of Consequences in Ethics", *International Journal of Ethics* 24 (1913):293-320, Lars Bergström, *The Alternatives and Consequences of Actions* (Stockholm: Almqvist & Wiksell, 1966), and J. Howard Sobel, "Utilitarianisms: Simple and General".

⁴ Note that if outcomes are probabilistic states of affairs, then theories of the good for teleological theories must also specify how the probabilities affect the goodness of the state of affairs. The usual assumption that the goodness value of a probabilistic state of affairs is equal to its expected goodness value is but one way of doing this. Other theories might, for example, specify that the goodness value of a probabilistic state of affairs is equal to the goodness value of the worst (or best) non-modal state of affairs that might be realized. If outcomes are possibilistic, but non-probabilistic, states of affairs, then equating goodness with objectively expected goodness is not even a possibility, and some other theory of the good is necessary.

⁵ Marcus Singer, "Actual Consequence Utilitarianism", *Mind* 86 (1977):67-77.

⁶ Singer, "Actual Consequence Utilitarianism", p.74.

⁷ Singer, "Further on Actual Consequence Utilitarianism", *Mind* 92 (1983):270-274, pp.272-274; and "Incoherence, Inconsistency, and Moral Theory: More on Actual Consequence Utilitarianism", *Southern Journal of Philosophy* 20 (1982):375-391, pp.377-379, pp.387-391. These two papers are replies to criticisms (noted below) by Jack Temkin, "Actual Consequence Utilitarianism: A Reply to Professor Singer", *Mind* 87 (1978):412-414, and by Loren E. Lomasky, "Is Actual Consequence Utilitarianism Incoherent?", *Southern Journal of Philosophy* 16 (1978):71-78, respectively.

⁸ Singer, "Actual Consequence Utilitarianism", pp.74-76; Singer, "Incoherence, Inconsistency, and Moral Theory", pp.381-384.

⁹ Both Temkin, "Actual Consequence Utilitarianism", and Lomasky, "Is Actual Consequence Utilitarianism Incoherent?" make this point.

¹⁰ Of course, *some* aspects of the future may be determined. Indeterminism does not rule out partial, local determinism, but that is of no help here.

¹¹ Singer, "Actual Consequence Utilitarianism", p.70, and Singer "Incoherence, Inconsistency, and Moral Theory", pp.376-377, p.381, and p.384.

¹² I here understand coherence as requiring not only logical consistency but also the absence of presupposition failure. Teleological theories that base the permissibility of actions on the goodness of what happens are incoherent because they falsely presuppose the existence of certain sorts of facts about what happens.

¹³ Donald Regan, *Utilitarianism and Co-operation* (Oxford: Clarendon Press, 1980).

¹⁴ Throughout I use 'joint action' in the sense of 'maximal joint action', i.e., in the sense that it includes an action component for every agent.

¹⁵ Regan, *Utilitarianism and Co-operation*, p.26. His matrix and probabilities are different, but the form of the argument is the same.

¹⁶ See Regan, *Utilitarianism and Co-operation*, pp.12-16, and n1., p.230.

¹⁷ Regan, *Utilitarianism and Co-operation*, p.18.

¹⁸ Note that, although like the incoherent version of utilitarianism correctly attacked by Singer, this version also makes the permissibility of actions depend on what happens

while or after they are performed, this version is not incoherent. This is because the only "happenings" that are relevant are the behaviour of the other agents at the time of action. There is no problem in determining the outcomes of unperformed actions on the assumption that the other agents behaved in certain ways.

¹⁹ Indeed, the given formulation of utilitarianism does not even agree with his official (and the most natural) definition of act utilitarianism (p.12) according to which act utilitarianism judges an action permissible just in case it has a maximally good outcome (e.g., consequence). The given formulation is nonetheless the one he uses in his examples, and the one he describes unofficially (e.g., pp.230-233).

²⁰ J. Howard Sobel makes this point (although in a different way and in a different context) in his critical notice of Regan's book, "Utilitarianism and Co-operation", *Dialogue* 24 (1985):137-152.

²¹ Regan, *Utilitarianism and Co-operation*, pp.54-55.

²² Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984) follows Regan and claims on p.54 that agent-neutral, individualistic, consequentialist theories (that is, theories that judge an action permissible just in case *its* outcome is *from a specified agent-neutral viewpoint* maximally good) cannot be directly collectively self-defeating. He claims, that is, that such theories are such that being a joint action each of the components of which satisfies the theory is compatible with being a joint action with a best outcome. Parfit makes it explicit that he takes the eventual behaviour of the other agents as given: "According to C [his standard, agent-neutral, individualistic, consequentialist theory], each of us should try to do what would the outcome the best, *given what others will actually do.*" (p.30, emphasis in text). Like the case with Regan, this claim is true of the "eventual behaviour" version of utilitarianism, but not of the most natural version.

²³ Regan, *Utilitarianism and Co-operation*, n.10, p.245.

²⁴ I have benefited from the critical comments on an earlier version of this paper made by Shelly Kagan and J. Howard Sobel.