

KRISTER BYKVIST

ALTERNATIVE ACTIONS AND THE SPIRIT OF  
CONSEQUENTIALISM \*

(Received 19 April 2000; received in revised version 13 April 2001)

**ABSTRACT.** The simple idea behind act-consequentialism is that we ought to choose the action whose outcome is better than that of any alternative action. In a recent issue of this journal, Erik Carlson has argued that given a reasonable interpretation of alternative actions this simple idea cannot be upheld but that the new theory he proposes nevertheless preserves the act-consequentialist spirit. My aim in this paper is to show that Carlson is wrong on both counts. His theory, contrary to his own intentions, is not an act-consequentialist theory. By building on a theory formulated by Holly Smith, I will show that the simple idea can be upheld. The new theory I will propose has all the merits of Carlson's theory without sharing its demerits.

1. INTRODUCTION

Consequentialism has been accused of many things. It has been accused of being too demanding, too permissive and too impartial, just to mention a few complaints. However, critics of consequentialism usually concede that it has at least one virtue: It is a reasonably clear and precise normative theory. It simply tells us to perform the action that will bring about the best consequences. More precisely, consequentialism in its maximizing guise is the conjunction of the following criteria.

*Maximizing Consequentialism (Max)*

An action is *obligatory* for (a person) *P* in (a situation) *S* iff its outcome would be better than the outcome of any alternative action for *P* in *S*.

---

\* Earlier versions of this paper were presented at *Filosofidagarna i Göteborg*, the Swedish National Conference in Philosophy, held at the University of Gothenburg, June 12, 1999, and to the audience at John Broome's seminar at the University of St Andrews, May 10, 1999.



An action is *right* for  $P$  in  $S$  iff its outcome would not be worse than the outcome of any alternative action for  $P$  in  $S$ .

An action is *wrong* for  $P$  in  $S$  iff it is not right for  $P$  in  $S$ .

This would be a precise characterization if only we knew how to select the alternative actions in a choice situation. This problem was not seriously investigated until the sixties, but since then a great number of answers have been suggested. In a recent issue of this journal, Erik Carlson presented a new proposal.<sup>1</sup> He argued that given a reasonable interpretation of alternative actions *Max* cannot be upheld but that the new theory he proposed nevertheless preserves the consequentialist spirit of *Max*. My aim in this paper is to show that Carlson is wrong on both counts. His theory, contrary to his own intentions, is not an act-consequentialist theory. By improving on a theory of alternative actions formulated by Holly Smith, I will show that *Max* can be upheld. The new theory I propose has all the merits of Carlson's theory without sharing its flaws.

## 2. PRELIMINARIES

Before I start presenting Carlson's theory I shall say something about the framework I intend to work with. I assume with Carlson that *counterfactual determinism* is true, i.e., that every action has a unique possible world associated to it – the world that would be realized were the action performed. I also assume, again with Carlson, that the outcome of an action is the possible world, or the possible future, that would be the case if the action were performed. These are exacting assumptions, especially the first one, but in the debate on alternative actions they are usually taken for granted and I will therefore not question them here.<sup>2</sup>

I assume that consequentialism deals with *particular* actions (e.g., Jane's waving her hand at noon, November 9, 2000), as opposed to *generic* actions (e.g., waving, kissing, or screaming). Furthermore, I take these particular actions to be abstract in the sense that we can meaningfully talk about *unperformed* actions. In this respect particular actions are similar to contingent states of affairs that need not obtain. Finally, I assume that there are

*compound* actions in the sense that one action can contain another action. An action  $a$  is *contained* in (a *part of*) another action  $a'$  iff  $a$  and  $a'$  are agent-identical, the time-interval of  $a$  is a proper or improper part of the time-interval of  $a'$ , and necessarily, if  $a'$  is performed then  $a$  is performed.<sup>3</sup> An example of a compound action is Jane's going to the bank at  $t$  and Jane's withdrawing money for Oxfam at  $t'$ . Again, this sketchy act-ontology is widely accepted by the participants in the debate.<sup>4</sup>

Concerning the notion of performability, I have no full-fledged theory to offer. I take it that normative theories, such as consequentialism, are primarily concerned with intentional actions. I therefore assume, again with Carlson, that our ability to act should be understood in terms of what we can do *intentionally*.<sup>5</sup> This means that an act is performable at a time only if the agent can at this time form an intention to do it. Since one cannot form an intention to do a certain action if one does not have the concepts necessary to grasp the action, an act is performable only if the agent can grasp the act. This will have consequences for our ability to perform acts that stretch into the future. For instance, when I was a young child I did not have the concept of being a philosopher. Thus no action that contains the action of becoming a philosopher was performable by my younger self.

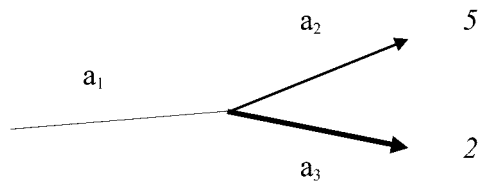
I do, however, assume that we sometimes can perform actions that extend into the future. It is, for instance, within my power at this very moment to type the letter 'J'. But it is also within my current power to type the name 'Jane'. But this is an action that extends into the future. So, we have to be careful to distinguish between the time of the power and the time of the action over which we have the power. These things are all uncontroversial. What is controversial is how this power over the future should be analysed. However, for my purposes, I do not need to give a full analysis.<sup>6</sup>

Finally, I shall say that an action is *immediately performable* by an agent at a time  $t$  iff the agent has an ability at  $t$  to perform the action and the action starts at  $t$  (or slightly later than  $t$ ).<sup>7</sup>

## 3. CARLSON'S THEORY

Carlson accepts the simple idea that the alternatives for an agent  $P$  in a situation  $S$  are all those actions that are performable by  $P$  in  $S$ . He claims that all other definitions will lead to difficulties. Given this definition he thinks he can no longer stick to *Max* because this would have the implication that some actions are obligatory only if they are not performed. To see this implication, assume the following choice situation.

*Case 1*



Action  $a_1$  and the two compound actions  $a_1$ -and- $a_2$  and  $a_1$ -and- $a_3$  are all performable by the agent, but  $a_1$ -and- $a_2$  is not performed in this situation. The thick line represents the fact that if the agent were to do  $a_1$ , he would do  $a_3$ . He would in no way be forced to choose  $a_3$ ; it is just that he would do  $a_3$  as a matter of fact. If we apply *Max* and assume that all performable actions are relevant alternatives, we get the result that  $a_1$ -and- $a_2$  is obligatory since this action has the best outcome. But note that  $a_1$ -and- $a_2$  would no longer be obligatory if the situation were such that  $a_2$  (rather than  $a_3$ ) would follow  $a_1$ . For in this counterfactual situation  $a_1$ -and- $a_2$  would be performed which means that  $a_1$  and  $a_1$ -and- $a_2$  would have the same outcome, i.e., the outcome with the value 5, and they would thus both be right. (Recall that the outcome of an action is the possible world, or the future, that would be the case if the action were performed.) Hence,  $a_1$ -and- $a_2$  is obligatory only if it is not performed.

This implication seems odd if we assume that a moral theory should guide actions. For how can a moral theory guide our actions if in order to decide what to do we have to know in advance what available actions we *will* perform? (I will return to the issue about action-guidingness in subsection 7a.)

To avoid this implication, Carlson assigns obligatoriness and rightness only to actions that *guarantee* an optimal outcome. An

action guarantees an optimal outcome in a certain situation just in case it would have an optimal outcome no matter what the agent did in this situation, i.e., no matter which of those actions available in the situation the agent would perform. The exact formulation of Carlson's theory requires some definitions.<sup>8</sup> An action is *optimal* for  $P$  in  $S$  iff its outcome is at least as good as that of any other alternative action for  $P$  in  $S$ . An action is *suboptimal* for  $P$  in  $S$  iff its outcome is worse than that of some other alternative action for  $P$  in  $S$ . An action is *invariably optimal* (*suboptimal*) for  $P$  in  $S$  iff it would have an optimal (suboptimal) outcome in  $S$  no matter what the agent would do in  $S$ . If in addition the action does not contain any invariably optimal actions, the action is a *minimally specific* invariably optimal action. Finally, if the action is unique in this respect it is a *unique* minimally specific invariably optimal action. In *Case 1*,  $a_1$ -and- $a_2$  is a unique minimally specific invariably optimal action, but  $a_1$  is not, since it would be suboptimal if  $a_1$ -and- $a_3$  were performed.

Carlson's theory can now be expressed as the conjunction of the following conditions.<sup>9</sup>

*K*

An action is obligatory for  $P$  in  $S$  iff it is a unique minimally specific invariably optimal action for  $P$  in  $S$ .

An action is right for  $P$  in  $S$  iff it is invariably optimal for  $P$  in  $S$ .

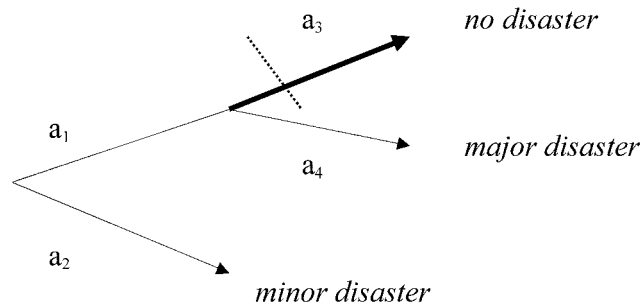
An action is wrong for  $P$  in  $S$  iff it is invariably suboptimal for  $P$  in  $S$ .

The invariability constraint is, of course, supposed to make normative status insensitive to facts about what the agent would do in the situation. In *Case 1*, *K* would, for instance, imply that  $a_1$ -and- $a_2$  is obligatory ( $a_1$ -and- $a_2$  is the least specific action that would have an optimal outcome no matter what the agent would do in the situation),  $a_1$ -and- $a_3$  is wrong ( $a_1$ -and- $a_3$  would have suboptimal consequences no matter what the agent would do), and  $a_1$  lacks normative status ( $a_1$  is neither obligatory, right, nor wrong since it is neither invariably optimal, nor invariably suboptimal), and that this holds no matter what actions the agent would perform in the situation. By demanding minimal specificity we avoid a shortage of obligations. There is seldom a uniquely optimal action in a situation since most actions have more specific versions that would not make

a difference in value. For instance, right now I can sign a check for Oxfam in black ink and mail it but I can also sign a check for Oxfam in blue ink and mail it.

Is  $K$  acceptable? As it stands, it will have counterintuitive normative implications. Consider this schematic example.

*Case 2*



Again, the thick line represents the fact that the agent would freely do  $a_3$  if he were to do  $a_1$ . The dotted line represents the fact that the compound act  $a_1$ -and- $a_3$  is not performable by the agent in his present situation (although  $a_3$  will be performable once he has reached the future choice point). So, the only way to avoid a disaster is to perform  $a_1$ . Consequently, it seems reasonable to say that the agent ought to do  $a_1$  (or at least that  $a_1$  and  $a_2$  should not share the same normative status). But  $K$  denies this since  $a_1$  is not invariably optimal.  $a_1$  would be suboptimal if  $a_1$ -and- $a_4$  were performed, for in that case  $a_1$  would have the same outcome as  $a_1$ -and- $a_4$ , i.e., a major disaster, and the consequences of  $a_2$  would be better (less bad).<sup>10</sup> Carlson would therefore have to say that  $a_1$  lacks normative status. Note also that  $a_2$  is not invariably suboptimal. It would have been optimal if the situation had been such that  $a_1$ -and- $a_4$  was performed. Hence  $a_2$  lacks normative status as well. The only action that has normative status is  $a_1$ -and- $a_4$ ; this action is invariably suboptimal and thus wrong.

One could, of course, wonder if this kind of case is possible. If I would do  $a_3$  were I to do  $a_1$  and I can do  $a_1$ , must we not also say that I can do  $a_1$ -and- $a_3$ ? Not if, following Carlson, we make the reasonable assumption that an act is performable at a time only if the agent can at this time form an intention to do it.<sup>11</sup> For

suppose that the agent at the original choice point lacks the concept of  $a_3$ . Imagine, for instance, that  $a_3$  is the act of formally deducing a theorem in predicate logic and that at the original choice point the agent lacks the concept of formal predicate logic. Then since he lacks the concept of  $a_3$ , he cannot at this time form the intention to do  $a_3$  or any act that includes  $a_3$ . Consequently,  $a_1$ -and- $a_3$  is not performable at this time.

However, Carlson (1995) formulates an alternative theory, called  $K^*$ , which would avoid the implication that only  $a_1$ -and- $a_4$  has normative status.<sup>12</sup> Roughly put,  $K^*$  agrees with  $K$  in prescribing an invariably optimal action in those situations in which there is such an action. However, in cases where there are no such actions, then  $K^*$  differs from  $K$  in prescribing actions that are merely optimal. So, in *Case 2*,  $K^*$  would tell the agent to do  $a_1$  since this action is optimal and there are no invariably optimal actions. Similarly,  $K^*$  permits the agent to perform merely optimal actions in cases where there are no invariably optimal actions. Finally,  $K$  and  $K^*$  provide the same criteria for wrongness. In the following I will stick to the simpler theory  $K$ . My subsequent criticism will not hinge on whether merely optimal actions are assigned normative status.

Carlson admits that  $K$  departs from traditional consequentialism in certain respects.<sup>13</sup>  $K$  entails that an action can be obligatory without having a uniquely optimal outcome (recall that an action can be obligatory even though it has a more specific version that has the same outcome), that it can have an optimal outcome without being right (recall that an act is right only if it is *invariably* optimal), and that it can have a suboptimal outcome without being wrong (recall that an act is wrong only if it is *invariably* suboptimal). Nevertheless he claims that  $K$  preserves the spirit of consequentialism. I agree that a theory can give up these conditions for normative status and still be true to the consequentialist spirit. The problem, however, is that  $K$  entails that the normative status of alternative actions is not exclusively determined by the values of their outcomes. This, I will argue, is at odds with consequentialism.

## 4. THE SPIRIT OF CONSEQUENTIALISM

The basic consequentialist intuition is that alternative acts should be judged exclusively by the values of their outcomes. The normative status of alternative actions – their being right, wrong or obligatory – depends solely on the values of their outcomes. In saying this, I assume act-consequentialism. This intuition does not underlie indirect versions of consequentialism since they detach normative status of alternative acts from their individual outcome-values and determine the normative status of an action indirectly by appeal to the values of the outcomes of things other than the alternative actions in the situation (e.g. rules or motives).

This act-consequentialist intuition is partly captured by the following supervenience condition.

*Consequentialist Supervenience (CS)*

If two actions available in the same choice situation would bring about outcomes with the same intrinsic value, then the actions have the same normative status in the situation.

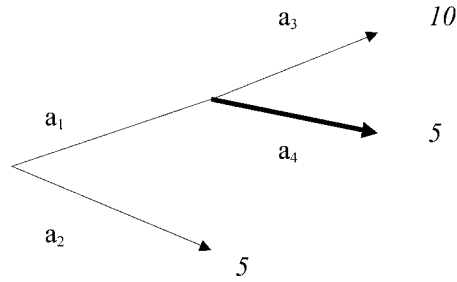
Surely, if nothing but outcome-value is supposed to matter for normative status, then two actions in the same situation cannot differ in normative status, if they have exactly the same outcome-value.<sup>14</sup>

Given the reasonable assumption that two equally good outcomes always occupy the same position in the ordering of all available outcomes, a maximizing consequentialism satisfies *CS*. The supervenience condition is also satisfied by non-maximizing forms of consequentialism, such as satisficing consequentialism, which tells us to choose the action whose outcome is “good enough”, and non-comparative consequentialism, which claims that the normative status of an action is wholly determined by the absolute value of its outcome, for instance, by saying that an action is right if its outcome is good, wrong if its outcome is bad.<sup>15</sup>

However, Carlson’s theory *K* does not satisfy *CS*. To see this, consider the following example.



## Case 3

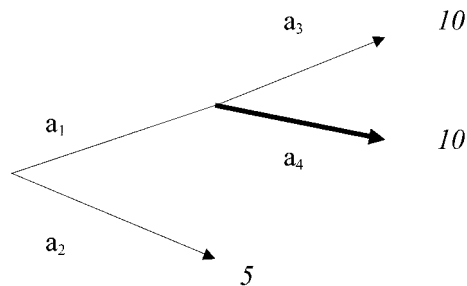


In this situation, if the agent were to perform  $a_1$ , he would later freely perform  $a_4$ . (This is represented by the thick arrow.) The outcome of  $a_1$  in this situation is thus identical with the outcome of  $a_1$ -and- $a_4$ .

If, as Carlson assumes, the alternative actions are all those actions that are performable by the agent, then *CS* dictates that  $a_1$ -and- $a_4$  and  $a_1$  should have the same normative status since they have exactly the same outcome-value. But Carlson's theory *K* implies that they have different normative status.  $a_1$ -and- $a_4$  is wrong since it would bring about a suboptimal outcome no matter what the agent were to do in the situation.  $a_1$  lacks normative status since it guarantees neither an optimal outcome, nor a suboptimal outcome:  $a_1$  would be optimal if the agent were to do  $a_3$ , but would be suboptimal were he to do  $a_4$ .

This example shows that *K* departs from the spirit of consequentialism by supposing that the normative status of an alternative action may depend on the value of the outcome it would bring about in a *counterfactual* situation in which the agent acted differently. However, this is not the only respect in which it departs from consequentialism. Consider the following example.

## Case 4



$K$  entails that  $a_1$  is obligatory in this case since it is a unique minimally specific invariably optimal action.  $a_1$ -and- $a_4$  is right but not obligatory since it contains an invariably optimal action. But since all performable actions are alternatives,  $CS$ , on the other hand, entails that  $a_1$  and  $a_1$ -and- $a_4$  share the same normative status. Here  $K$  departs from the spirit of consequentialism by assuming that whether an action has a certain normative status may depend on the *internal relations* between the alternative actions, i.e., whether one optimal action includes another.

So, I think it is clear that  $K$  does not preserve the act-consequentialist spirit. It entails that the normative status of an alternative action does not depend solely on the value of its outcome and the value of the outcomes of its alternatives. This is not to say that the theory cannot be seen as some kind of consequentialism. It is just that it does not qualify as *act*-consequentialism. Rather it should be seen as an indirect form of consequentialism since it detaches the normative status of the alternative actions from their individual outcome-values.

One could, however, argue that this means that my criticism of Carlson loses its force. I have shown that Carlson's theory  $K$  does not capture the act-consequentialist spirit. But so what? Perhaps he only wanted to capture the spirit of an indirect consequentialism.

I do think it is clear that Carlson's intention was to formulate an act-consequentialist theory. In the beginning of his (1995) Carlson assumes that consequentialism includes a principle dubbed "axiologism" ( $AX$ ) and which is formulated thus:

$AX$  The moral status of an action  $a$  is wholly determined by the absolute or relative intrinsic value of the outcome of  $a$ .<sup>16</sup>

The relative intrinsic value of an action's outcome is defined as the value the outcome has as compared to the values of the outcomes of alternatives to  $a$ .  $AX$  must be taken to entail that if two actions in the same situation have the same absolute and relative intrinsic value then they must have the same normative status. But this means that  $AX$  entails  $CS$ , since if two outcomes in a certain situation have the same intrinsic value then, obviously, they must have the same absolute and relative intrinsic value. So, something has to go: either

act-consequentialism or Carlson's theory *K*. I shall show that there is no need to abandon act-consequentialism in order to solve the problems of alternative actions. More specifically, I shall show that there is a reasonable theory that satisfies *Max* and preserves the consequentialist spirit.

## 5. MAXIMALISM

Since my theory draws heavily on a consequentialist theory proposed by Holly Smith, I will start by presenting her views.<sup>17</sup> Roughly put, her theory ascribes normative status only to actions that are maximal in the sense of not being proper parts of anything the agent can do. In particular, the theory claims that the agent ought to perform the maximal action that would lead to the best outcome. To render this theory more precise, we need to introduce the concept of a maximal action. Using the notions of 'parthood' and 'immediate performability' defined in section 2, we can say that an action *a* is a *maximal* action for *P* in *S* iff *a* is immediately performable by *P* in *S* but not contained in any other action immediately performable by *P* in *S*.

Smith's theory amounts to the claim that the alternative set for *P* in *S* consists of all maximal actions for *P* in *S* and that *Max* should be applied to these actions. So, an act is obligatory iff it is a maximal action whose outcome is better than that of every other maximal action, right iff it is a maximal action whose outcome is at least as good as that of every other maximal action, and wrong iff it is not right. It is clear that Smith's theory preserves the spirit of consequentialism since it satisfies *Max*. However, as it stands it will have some counterintuitive consequences. Go back to *Case 2*. Here the maximal actions are *a*<sub>1</sub>-and-*a*<sub>4</sub>, and *a*<sub>2</sub>, and both will have disastrous consequences. (Recall that *a*<sub>1</sub>-and-*a*<sub>3</sub> is not performable.) Smith's theory would therefore claim that *a*<sub>2</sub> is obligatory since, of all the maximal actions, this action would have the best consequences (it will only lead to a minor disaster). But this prescription is absurd since there is a way to avoid a disaster, namely, to perform the non-maximal action *a*<sub>1</sub>.

I shall show that there is a way to modify Smith's theory so that this implication is avoided and that this can be done without

abandoning act-consequentialism. Roughly put, my idea is that the alternative set for  $P$  in  $S$  consists of all maximal actions for  $P$  in  $S$  plus every non-maximal performable action  $a$  such that if  $P$  were to perform  $a$ , then the only distinct actions he would perform (if any) would be actions that in conjunction with  $a$  constitute actions not immediately performable by  $P$  in  $S$ . More precisely, the condition on non-maximal actions amounts to the following:

For any action  $b$  not contained in  $a$ , if  $b$  would be performed were  $P$  to perform  $a$  in  $S$ , then  $a$ -and- $b$  is not immediately performable by  $P$  in  $S$ .

Note that all maximal actions satisfy this condition. To see this, assume that  $a$  is a maximal action, i.e.,  $a$  is not a proper part of any action immediately performable by the agent. Assume for reductio that there is an action  $b$  not contained in  $a$  such that  $b$  would be performed were the agent to perform  $a$  and that  $a$ -and- $b$  is immediately performable by the agent. Then  $a$  is not a maximal action since it is a proper part of an immediately performable action:  $a$ -and- $b$ . But then we have a contradiction.

This means that we can simplify and say that the alternative set for  $P$  in  $S$  consists of every action  $a$  such that (1)  $a$  is immediately performable by  $P$  in  $S$ , and (2) for any action  $b$  not contained in  $a$ , if  $b$  would be performed were  $P$  to perform  $a$  in  $S$ , then  $a$ -and- $b$  is not immediately performable by  $P$  in  $S$ . We could call the actions that satisfy (1) and (2) *quasi-maximal* actions. My proposal then is that *Max* should be applied only to quasi-maximal actions. In particular, it tells the agent to perform the quasi-maximal action that would have the best consequences.

Modified in this way, maximalism will not have any counter-intuitive consequences for *Case 2*. The quasi-maximal actions in this situation are  $a_1$ ,  $a_1$ -and- $a_4$ , and  $a_2$ .<sup>18</sup> Of these actions  $a_1$  has the best outcome. Hence,  $a_1$  ought to be performed, whereas  $a_1$ -and- $a_4$  and  $a_2$  are wrong.

The theory gives the correct prescriptions not just for *Case 2* but also for cases 1, 3 and 4. Consider *Case 1*. Here the quasi-maximal actions are  $a_1$ -and- $a_2$  and  $a_1$ -and- $a_3$ . (These actions are maximal and thus quasi-maximal.) Since  $a_1$ -and- $a_2$  has a better outcome, it ought to be done. In *Case 3*,  $a_1$ -and- $a_3$ ,  $a_1$ -and- $a_4$ , and  $a_2$  are quasi-

maximal simply because they are all maximal. Here  $a_1$ -and- $a_3$  is obligatory and  $a_2$  and  $a_1$ -and- $a_4$  are both wrong.  $a_1$  lacks normative status since it is not a quasi-maximal action; it would be followed by  $a_4$ , and  $a_1$ -and- $a_4$  is immediately performable. Finally, in *Case 4*,  $a_1$ -and- $a_3$ ,  $a_1$ -and- $a_4$ , and  $a_2$  are quasi-maximal since they are all maximal. Here  $a_1$ -and- $a_3$  and  $a_1$ -and- $a_4$  are right,  $a_2$  is wrong, and  $a_1$  again lacks normative status.

One could object that it is not reasonable to say that actions that are proper parts of quasi-maximal actions lack normative status. Surely,  $a_1$  in *Case 3* has some normative significance, since you have to perform this action *in order to* do what you ought to do. I think the right thing to say is that we have to distinguish between what an agent ought to do in the *non-derivative* sense and what he ought to do in the *derivative prerequisite* sense. Normative theories and constraints on normative theories such as *CS* are primarily concerned with what is obligatory in the non-derivative sense. Actions have derivative normative status in virtue of being parts, and thus prerequisites, of actions that have normative status in the non-derivative sense. The prerequisite sense of ought could then be defined thus.<sup>19</sup>

An action *ought* to be done in the prerequisite sense iff it is contained in every right action.

This definition implies that actions that are parts of obligatory actions are obligatory in the prerequisite sense. For if an action is obligatory in a situation then it is the only right action in the situation. Even in situations where no action is obligatory the definition will sometimes entail that some action(s) ought to be done in the prerequisite sense. Take, for instance, *Case 4*. Here we have two maximal actions with optimal outcomes and thus both are right but neither is obligatory. But since  $a_1$  is part of both actions it will be obligatory in the prerequisite sense. This seems reasonable, since in this case the agent has to perform  $a_1$  in order to secure an optimal outcome.

In contrast to *K*, my theory satisfies *CS* since it satisfies *Max*. So, my theory is in line with the basic consequentialist intuition. Moreover, my theory shares with *K* three important merits. First, it satisfies *Normative Harmony*, i.e., the principle that if *P* in *S* does

what he ought to do, then  $P$  does not in  $S$  do anything wrong.<sup>20</sup> Second, it satisfies *Compatibility of Situation-Identical Obligations*, i.e., the principle that if two actions are mutually incompatible for  $P$  in  $S$ , then they are not both obligatory for  $P$  in  $S$ .<sup>21</sup> Third, my theory never requires an act that would bring about an outcome that is suboptimal in the sense of being worse than the outcome of some other *performable* action. Smith's theory, on the other hand, will sometimes – in *Case 2*, for instance – require acts with suboptimal outcomes. (In *Case 2*, Smith's theory requires  $a_2$  even though  $a_1$  has better consequences.)

## 6. ACTUALISM AND POSSIBILISM

Like Carlson's theory my theory occupies an intermediate position in the debate on actualism and possibilism, i.e., the debate about whether an agent's own simultaneous or future actions are relevant for determining what the agent should do in a certain situation. Whereas extreme actualism claims that present and future actions are always relevant and extreme possibilism claims that they are never relevant, my theory claims that they are sometimes relevant. They are relevant only if they are beyond the control of the agent, i.e., not included in an action that is performable by the agent.

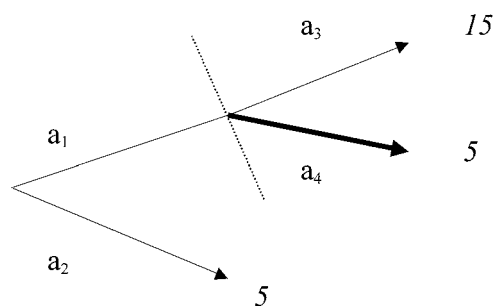
I agree with Carlson that extreme actualism and extreme possibilism are not plausible. As Carlson observes, one major problem with actualism is that it is too tolerant of moral imperfection.<sup>22</sup> Suppose I am a selfish miser and therefore not keen on sending money to charity. Because of my disposition it is true that if I were to go to the bank, I would not withdraw money for Oxfam, even though I could. Going to the bank and failing to withdraw money would be worse than staying home and saving the money, since if I stayed home I would not have to bother getting to the bank in rain and heavy traffic. But going to the bank and withdrawing money for Oxfam would still have the best consequences, since that would help relieve the suffering of others. Now, according to actualism, I ought to stay home, since this has better consequences than going to the bank.<sup>23</sup> This is counterintuitive. Surely, my own selfishness cannot get me off the moral hook so easily.<sup>24</sup>

Even though I have nothing to add to Carlson's criticism of actualism, I think there is something more that can be said against possibilism: If you want to stick to act-consequentialism, a certain kind of possibilism is no option since it violates *CS*. The kind of possibilism I have in mind is the life-oriented approach according to which what an agent should do is determined by the values of the outcomes of the *lives* still open to him at the time of decision. More specifically, this theory states that an action is obligatory if it is part of a life with optimal consequences.

A possible life for an agent, roughly put, is a compound action, lasting throughout the remainder of the agent's life, every part of which is such that the agent would be able to perform it, if he were to perform the earlier parts.<sup>25</sup> However, a possible life is usually not a performable action. Even if each part of a possible life is performable at its respective time, it is not true that a possible life *as a whole* is performable at its first moment by an agent since she may not then be able to grasp some of the actions contained in the life. If she cannot grasp an action as a whole then she cannot form an intention to perform it and, in the morally relevant sense, she cannot perform it. Recall the example from section 2 about my younger self who lacked the concept of being a philosopher. Since I lacked this concept, it seems plausible to say that possible lives containing the action of becoming a philosopher were not performable by my younger self.

To see that my formulation of consequentialism excludes the life-oriented theory, consider the following example.

*Case 5*



In this situation the agent can end his life by choosing  $a_2$ . He can also continue to live (by choosing  $a_1$ ) and make one more choice before he dies. However, the conjunction of  $a_1$  with either of its successors is not immediately performable by the agent in his present situation (although each successor *will* be performable once he has reached the future choice point). Perhaps he lacks the concept needed to grasp the compound actions  $a_1$ -and- $a_3$  and  $a_1$ -and- $a_4$ . If he were to choose  $a_1$ , he would as a matter of fact do  $a_4$  in the next choice situation.

Since  $a_1$  and  $a_2$  have outcomes with exactly the same value, CS implies that they must share the same normative status as well. But a defender of the life-oriented approach might object. He could argue that  $a_2$  is wrong since it is a suboptimal life whereas  $a_1$  is obligatory since it is *part* of an optimal life. In general, he could claim that we should start on the best life that is still open to us.<sup>26</sup>

That the life-oriented approach is at odds with act-consequentialism is perhaps not surprising. There is a striking similarity between this approach and an indirect theory such as rule-consequentialism. The former claims that the normative status of an action depends on what possible *life* it is a part of and the outcome-value of this life as compared to the outcome-values of alternative lives. The latter claims that the normative status of an action depends on what rule it can be subsumed under and the outcome-value of this rule (or rather of the general acceptance of this rule) as compared to the outcome-values of alternative rules. In both cases, the normative status of an action is determined *indirectly* via the outcome-values of things other than the actions performable by the agent.

Note that this feature of life-oriented theories makes for some troublesome implications. For starting on the best life still open to us might lead to some misery that could have been prevented had we performed another available option. A proponent of the life-oriented approach cannot be moved by considerations like this. In this respect, too, the life-oriented approach is similar to rule-consequentialism that may tell us to adhere to a rule even if breaking it would prevent some misery. In contrast, the act-consequentialist can still claim that starting on the best life (or adhering to a certain rule) is wrong if doing something else would have better consequences.



## 7. OBJECTIONS

(a) *Normative status depends on facts about what the agent will do.* As I mentioned in the beginning, what motivates Carlson to invoke the concept of an invariably optimal (suboptimal) action is the idea that the normative status of an action should not depend on what actions the agent would perform in the choice situation. More exactly, Carlson thinks that the following invariance condition should be satisfied.

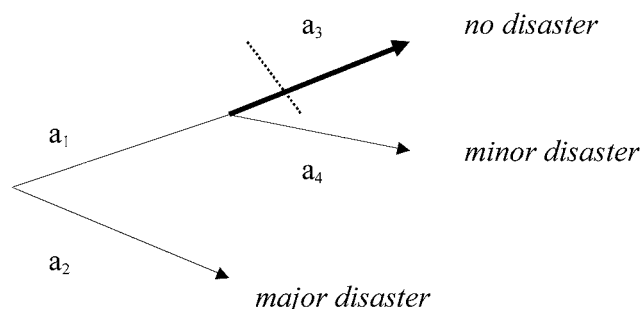
*Normative Invariance*

An action's normative status does not depend on whether or not it is performed.<sup>27</sup>

What is the intuitive appeal of this principle? Carlson claims that even though our main aim is to formulate objective criteria of rightness and not decision-making procedures, it seems reasonable to demand that normative theories should in principle be action-guiding for ideal agents – agents with full knowledge of all morally relevant facts in the situation.<sup>28</sup> Thus it should be possible for an ideal agent to deliberate and make her decision on the basis of the normative theory in question. But if the agent needs to know what he will actually do in *S* in order to know what he should do in *S* then he cannot deliberate about what to do. You cannot deliberate about whether to do *a* if you believe with certainty that you will do *a*. Of course, you may still be able to judge the action in question; you can check whether it is in accordance with a certain criterion of rightness – whether it maximizes happiness, for instance. But to make a decision to perform or not to perform a certain action is not to judge the action; it is to form an intention to perform or not to perform the action.

To see that my theory violates *NI*, consider *Case 6*. This situation is exactly like *Case 2* except that *a*<sub>2</sub> and *a*<sub>1-and-a</sub><sub>4</sub> have traded outcomes.

## Case 6



In this situation  $a_1$ -and- $a_4$  is wrong according to my theory, since there is another quasi-maximal action,  $a_1$ , that would have better consequences (no disaster). However, if  $a_1$ -and- $a_4$  were performed, it would be right since in that case  $a_1$  would no longer be a quasi-maximal action and hence no alternative. The reason  $a_1$  would no longer be a quasi-maximal action is that it would be followed by  $a_4$ , and  $a_1$ -and- $a_4$  is immediately performable. The only quasi-maximal alternative to  $a_1$ -and- $a_4$  would be  $a_2$ , which would have worse consequences (a major disaster).

But is this a good objection to my theory? Note that this situation is special in the sense that both the repertoire of alternative actions and the associated outcomes would change if  $a_1$ -and- $a_4$  were performed. If  $a_1$ -and- $a_4$  were performed, then  $a_1$  would no longer be an alternative action. The agent would face a choice between two disasters, one major and one minor. Why should we assume that the normative status of an action should be fixed when we vary the alternatives and their outcomes? After all, the underlying idea of a comparative consequentialism like my theory is precisely that the normative status of an action is determined by the value of its outcome as compared to the values of the outcomes of its alternatives.

This, however, does not mean that my theory is completely useless as an action-guide for ideal agents. In all cases where the agent faces only maximal options, my theory will not violate *NI* and thus the agent can know what he ought to do without knowing what he will do. In cases where the theory violates *NI* and the agent must know what he will do in order to know what he ought to do, my theory can still guide the agent. Even if it is conceded that deciding

to perform an action is logically incompatible with believing that one will do this action, this does not in itself imply that if one believes that one will perform an action then one is *unable* to decide to perform it. The only thing that is implied is that one *will* not decide to do the action.<sup>29</sup> To show that one cannot decide requires the additional assumption that the ideal agent's beliefs about what he will do are unavoidable for him. But I doubt that this assumption is reasonable. True, it might be reasonable to say that latent beliefs cannot change at will but, surely, it is only manifest beliefs that crowd out decisions. Why should there be a problem for me to decide to do a certain action if my belief that I will do it is not present to my consciousness? Since it seems strange to assume that an ideal agent's manifest beliefs about what he will do are unavoidable for him what we can show at most is that ideal agents who subscribe to theories that violate *NI* will not be guided by these theories.

In any case, Carlson's own theory *K* is not a better guide to action. It, too, will violate *NI*.<sup>30</sup> Consider *Case 6* again. According to *K*, *a<sub>1</sub>-and-a<sub>4</sub>* is not right in this situation since it is not optimal – *a<sub>1</sub>* has a better outcome. However, if *a<sub>1</sub>-and-a<sub>4</sub>* were performed then it would be invariably optimal and hence right. In this case, *a<sub>1</sub>-and-a<sub>4</sub>* and *a<sub>1</sub>* would have the same outcome and *a<sub>1</sub>-and-a<sub>4</sub>* would be optimal no matter what the agent were to do.

So, I am very doubtful whether *NI* is such a compelling constraint at least on the present reading. However, I do think there is one reading of the condition which is compelling. If we keep everything in two choice situations fixed except the actuality of a certain action, it seems reasonable to say that the normative status of this action cannot vary. More exactly, on this reading the condition states that if two situations involve the same alternative actions and the same associated outcomes but differ in that a certain action is performed in one situation but not in the other, then the normative status of that action does not vary from one situation to the other. This invariance condition is simply part of the consequentialist spirit. If acts should be assessed exclusively by the values of their outcomes, then the mere actuality of actions cannot make a normative difference. No wonder then that act-consequentialists are tempted to accept *NI*!

If this is the most reasonable interpretation of *NI*, then my theory will not have any problems with this constraint. Obviously,

if two situations involve exactly the same alternative actions and associated outcomes, then my theory will not judge the alternative actions from the two situations differently.

(b) *There are no maximal actions.* My theory will have a restricted use if every immediately performable action is contained in some other immediately performable action. For then there are no maximal action, and we can only assign normative status to merely quasi-maximal actions, i.e., actions such as  $a_1$  in cases 2 and 6. One could argue against the existence of maximal actions in the following way. Surely any adequate act-ontology must count actions as well as *inactions* (omissions). It seem natural to suppose that we can form act-compounds of actions and inactions. But then we can construct the following list of act-compounds:

$S$ 's raising his hand at  $t$  and not doing  $a_1$   
 $S$ 's raising his hand at  $t$  and not doing  $a_1$  and not doing  $a_2$   
 $S$ 's raising his hand at  $t$  and not doing  $a_1$  and not doing  $a_2$   
and not doing  $a_3$   
:                                       :

Since whatever I do I will always fail to perform an infinite number of actions this list will contain an infinite number of distinct actions, each of which is entailed by the subsequent ones. Hence we will not have any maximal actions.

But here we must be careful to distinguish inactions from *non-actions*. It is clear that whatever we do there is an infinite number of actions we fail to do in the sense that it is false to say that we perform any of them. But this does not entail that we intentionally perform an infinite number of omissions. In general, from the fact that it is false that I intentionally perform  $a$ , it does not follow that it is true that I intentionally refrain from doing  $a$ .<sup>31</sup> Suppose, for instance, that right now a child is drowning 1000 miles away from where I am, but I have no idea that this is happening. In this case, it would be absurd to say that I intentionally omit to save the child even though it is false that I intentionally save the child.

## 8. CONCLUSIONS

I have argued that there is no need to abandon the simple maximizing version of consequentialism (*Max*) if we identify the relevant alternatives with quasi-maximal actions. Furthermore, whereas Carlson's departure from *Max* is also a departure from act-consequentialism, my quasi-maximalist theory clearly captures the spirit of act-consequentialism.

It is true that my theory entails that in certain situations what we ought to do will depend on what we will do in the choice situation, but so will Carlson's theory. So, in this respect they are equal. In any case, theories that have this implication can still be used as guides for action.

Finally, if we carefully distinguish between non-derivative and derivative oughts, my theory will not have any normative implications that are implausible from a consequentialist perspective. Of course, this does not show that, *all things considered*, my version of consequentialism is preferable to *any* other moral theory. As I stated in the beginning, there are numerous other problems with consequentialism. I have shown, however, that the consequentialist can at least solve the problem of defining the alternative actions in a choice situation.

## ACKNOWLEDGEMENTS

For helpful comments, I would like to thank an anonymous referee for *Philosophical Studies*, Lars Bergström, John Broome, Garrett Cullity, Sven Danielsson, Kent Hurtig, Magnus Jiborn, Wlodek Rabinowicz, Michael Zimmerman, and, most particularly, Erik Carlson. It was Erik Carlson who first introduced me to the problem of alternative actions. Much of what is argued in this paper draws heavily on his work on alternative actions. Without his thorough treatment of the problem I would not have been able to propose a new solution. I am grateful for financial support provided by the Bank of Sweden Tercentenary Foundation.

## NOTES

<sup>1</sup> Carlson (1999).

<sup>2</sup> For statements of these assumptions, see, for instance, Carlson (1999), p. 256, n. 8; Carlson (1995), pp. 12, 70; Feldman (1997), pp. 29–31; Goldman (1976), p. 452; and Zimmerman (1996), pp. 25–31. If counterfactual determinism is rejected, then not every action has a unique outcome. Instead, each action should be assigned a *risky prospect*, a probability distribution over some set of states of affairs, each member of which has a certain probability obtaining given that the action is performed. A probabilistic consequentialism would then say that actions should be assessed by the values of their associated risky prospects. The theory of alternative actions I propose in this paper will, with suitable adjustments, work in a probabilistic framework, but space considerations prevent me from elaborating on this point.

<sup>3</sup> This notion of parthood is from Prawitz (1968), p. 80.

<sup>4</sup> One notable exception is Tännsjö (1978), who argues that particular actions are *concrete*, i.e., spatiotemporal entities.

<sup>5</sup> This sense of ‘can’ is the sense that figures in the principle “ought implies can”.

<sup>6</sup> My sketchy remarks on performability leaves it open whether or not the ability to act should be given a conditional analysis in terms of what I would do if I wanted or intended to do an action. Nothing in my discussion hangs on this, however. For a defence of a conditional analysis, see Carlson (1995), Chapter 5.

<sup>7</sup> The notion of immediate performability is found in Carlson (1995), p. 77.

<sup>8</sup> See Carlson (1999), p. 258, and Carlson (1995), pp. 102–103.

<sup>9</sup> Carlson (1999), p. 258. See also, Carlson (1995), pp. 99–109.

<sup>10</sup> Recall that I assume here that the outcome of an action is the possible world, or possible future, that would be actual if the action were performed.

<sup>11</sup> See Carlson (1995), pp. 81–82.

<sup>12</sup> Carlson (1995), pp. 107–109.

<sup>13</sup> Carlson (1999), p. 260.

<sup>14</sup> Of course, consequentialism implies something more. It also yields an invariance constraint that holds *across* different situations. More precisely, it implies that if two actions *a* and *b* from two different choice situations have outcomes with the same value and the alternative actions to *a* and *b* can be paired off without remainder with respect to the value of the outcome they would bring about, then the actions have the same normative status. For more on this condition, see Bykvist (1999).

<sup>15</sup> “Good enough” can be defined in different ways. On the absolute approach, an action is right if the value of its outcome is above a certain absolute level. On the comparative approach, an action is right if the value of its outcome is above a certain percentage of the value of the best outcome available. For more on these kinds of satisficing theories, see Carlson (1995), pp. 13–19, and Slote (1985).

<sup>16</sup> Carlson (1995), p. 9.

<sup>17</sup> Goldman (1978). (She has changed her name from Goldman to Smith since

the publication of this article.) My presentation of her theory differs slightly from her own presentation, but nothing crucial hangs on this.

<sup>18</sup>  $a_1$ -and- $a_4$  and  $a_2$  are quasi-maximal since they are maximal actions, and, as shown above, any maximal action is a quasi-maximal action.  $a_1$  is not a maximal action since it is contained in  $a_1$ -and- $a_4$ . But  $a_1$  is quasi-maximal, because it is immediately performable and such that if it were performed, the only distinct actions that would follow would be actions that in conjunction with  $a_1$  constitute actions not immediately performable. (Remember that even though  $a_1$  would be followed by  $a_3$ ,  $a_1$ -and- $a_3$  is not immediately performable.)

<sup>19</sup> For a similar notion of prerequisite normative status, see Carlson (1995), pp. 142–144, and Goldman (1978), pp. 190–191.

<sup>20</sup> This principle is found in Carlson (1999), p. 256. See also Carlson (1995), p. 103.

<sup>21</sup> This principle is found in Carlson (1999), p. 261. See also Carlson (1995), p. 95.

<sup>22</sup> Carlson (1999), p. 261. My example is just a variation of an example he gives.

<sup>23</sup> I assume here the actualism advocated by Sobel (1976, 1982). This theory identifies alternative actions with *minimal* actions, i.e., actions that once begun cannot be stopped short of completion, and tells the agent to perform the minimal action that has the best consequences. A similar theory, intended to be incorporated into a probabilistic consequentialism, was suggested by an anonymous referee for *Philosophical Studies*.

<sup>24</sup> Another problem with actualism is that it tells me to treat my future action (withdrawing money for Oxfam) on a par with the future actions of other people. But this is to neglect the fact that my future action is within my current power. It should be stressed that this is counterintuitive only so long as it is clearly within my current power to go to the bank *and* withdraw money for Oxfam. If my selfishness is strong enough to constitute a psychological compulsion, then it is less counterintuitive to say that I ought not to go to the bank. But in this case it is no longer clear that it is within my current power to withdraw money from Oxfam. My future action should therefore be treated as the future action of another person, i.e., as an outcome of going to the bank.

<sup>25</sup> See Carlson (1995), p. 88.

<sup>26</sup> This comes close to the “neo-utilitarian” theory advocated in Feldman (1986). Possibilism is also defended in Zimmerman (1996), Chapter 6.

<sup>27</sup> Carlson (1999), p. 256. He also thinks that a moral theory should satisfy this stronger principle:

If  $a$  is performable by  $P$  in  $S$ , then the normative status of  $a$  does not depend on what action(s)  $P$  performs in  $S$ .

For more on these invariance principles, see Carlson (1995), pp. 100–102.

<sup>28</sup> Carlson (1999), pp. 256–257. See also Carlson (1995), pp. 100–102.

<sup>29</sup> It is a modal fallacy to reason from ‘it is impossible that both  $p$  and  $q$ ’ to ‘if  $p$ , then it is impossible that  $q$ ’.

<sup>30</sup> This is conceded in Carlson (1999), p. 267, footnote 13. However, he claims

that this holds only for “a rather special type of case” and that *MI* is less compelling in this type of case. If he is right about this then my theory cannot be faulted because of its violation of *NI* in *Case 6*.

<sup>31</sup> Remember that normative theories are primarily concerned with intentional actions.

#### REFERENCES

- Bykrist (1999): ‘Normative Supervenience and Consequentialism’, manuscript.
- Carlson, E. (1995): *Consequentialism Reconsidered*, Dordrecht: Kluwer.
- Carlson, E. (1999): ‘Consequentialism, Alternatives, and Actualism’, *Philosophical Studies* 96(3), 253–268.
- Feldman, F. (1986): *Doing the Best We Can*, Dordrecht: D. Reidel Publ. Company.
- Feldman, F. (1997): *Utilitarianism, Hedonism, and Desert*, Cambridge: Cambridge University Press, Cambridge.
- Goldman, H. (1976): ‘Dated Rightness and Moral Imperfection’, *The Philosophical Review* 85, 449–487.
- Goldman, H. (1978): ‘Doing the Best One Can’, in A.I. Goldman and J. Kim (eds.), *Values and Morals*, pp. 185–214.
- Prawitz, D. (1968): ‘A Discussion Note on Utilitarianism’, *Theoria* 34, 76–84.
- Slote, M. (1985): *Common-Sense Morality and Consequentialism*, London: Routledge and Kegan.
- Sobel, H. (1976): ‘Utilitarianism and Past and Future Mistakes’, *Noûs* 10, 195–219.
- Sobel, H. (1982): ‘Utilitarian Principles for Imperfect Agents’, *Theoria* 48, 113–126.
- Tännsjö, T. (1978): ‘The Morality of Abstract Entities’, *Theoria* 44, 1–18.
- Zimmerman, M.J. (1996): *The Concept of Moral Obligation*, Cambridge: Cambridge University Press.

*Cambridge, UK*

*E-mail: krister.bykvist@filosofi.uu.se*