

HOLLY S. GOLDMAN

DAVID LYONS ON UTILITARIAN GENERALIZATION

(Received 23 April, 1973)

In *Forms and Limits of Utilitarianism*, David Lyons argues that utilitarian generalization does not succeed in circumventing the classical objections confronting act utilitarianism.¹ When properly understood, he claims, utilitarian generalization is extensionally equivalent to act utilitarianism: the two theories yield identical prescriptions when applied to the same cases. This claim, if correct, undermines much of utilitarian generalization's appeal, for anyone who rejects act utilitarianism on the ground that it yields counterintuitive prescriptions will find little merit in a theory which yields the very same prescriptions.²

In this paper, I shall argue that Lyons' claim is incorrect: utilitarian generalization, as he interprets it, is not extensionally equivalent to act utilitarianism. This conclusion gives proponents of generalization little cause for celebration, since Lyons' version of utilitarian generalization yields, if anything, even less plausible prescriptions than act utilitarianism. However, once the nature of Lyons' version of the theory is clearly recognized, it is easy to avoid this difficulty by formulating a closely related version. Unfortunately the resulting theory is also unsatisfactory, for it proves incapable of yielding any prescriptions at all when applied to a wide range of morally significant cases. Thus neither of these versions of utilitarian generalization is an acceptable moral theory. However, since Lyons is wrong in thinking that his version of utilitarian generalization is the only correct one, the door remains open for some possibly more successful version of the theory to be proposed.

1. LYONS' VERSION OF UTILITARIAN GENERALIZATION

The most plausible form of utilitarian generalization may be formulated in the following manner:

(UG) An action is right if and only if the consequences of everyone's

performing that sort of action would be at least as good as the consequences of everyone's performing any alternative sort of action.³

The analogous form of act utilitarianism may be stated similarly:

(AU) An action is right if and only if the consequences of that action would be at least as good as the consequences of any alternative action.

It is clear that no prescriptions may be derived from UG until some interpretation is given to two of its central terms, namely 'everyone' and 'that sort of action'.⁴ Both terms are ambiguous: 'everyone' could be taken as referring to every moral agent, or everyone involved in a certain practice, or everyone who has the opportunity to perform the act in question, while 'that sort of action' might mean, for example, 'voting', or 'voting for Candidate A', or 'voting for Candidate A while the majority votes for Candidate B'. Since the consequences of 'everyone's performing that sort of action' depend on which of these alternatives is chosen, the interpretation of these terms determines the prescriptions generated by UG.

Lyons suggests that we collapse the two problems into one. According to his proposal, the term 'everyone' should be interpreted to mean "that class of persons, each of whom will have occasion to do the sort of thing specified, to each of whom such a course of action is or will be a practical possibility." (p. 31) Then, if aspects of the agent's identity turn out to be significant, we may simply include them in an expanded description of the action. The important issue thus becomes the problem of determining which description of an action is the relevant one – which description must hold true of other actions if they are to qualify as 'that sort of action'. According to Lyons, this issue cannot be resolved simply by determining what features of an action are, generally speaking, morally relevant. What we need is an account of those features of an action which are morally relevant to a *particular* moral theory, namely UG. We need a 'criterion of relevance' that enables us to pick out which properties of an action must be included in its description for purposes of applying UG. This criterion, moreover, must be implied by the theory itself. (p. 34)

Lyons states that the criterion of relevance appropriate to UG is one

according to which only the 'consequentially significant' properties of actions are relevant for the application of UG. He defines a 'consequentially significant' property of an action as one in virtue of which that action causes some event having either utility or disutility. (p. 60) This position is argued for in the following passage:

Pure teleological principles are concerned, ultimately, only with the values of consequences of actions. Thus, when pure teleological principles are applied, particular actions may be viewed only with respect to their teleologically significant properties, that is, those properties *in virtue of which* actions produce utilities or disutilities. The forms of utilitarian generalization are pure teleological principles. Therefore, the only legitimate candidates for inclusion in the description of an action for the application of a form of utilitarian generalization are... consequentially significant properties, that is, those properties in virtue of which actions have effects. (p. 57)

Thus, for example, if we apply UG to the traditional case of violating gas rationing restrictions during the war, the appropriate description of the action would include reference to the property 'occurring when it is necessary to conserve gas for the war effort'. It would *not* include reference to the property 'occurring on a Tuesday', for this property has no causal effects. (p. 57) Nor would it include reference to such properties as 'producing – 10 utiles', for this property *refers to* but does not *cause* the effect in question. In particular, Lyons' criterion implies that frequently the relevant description of an action must contain reference to the activities of other agents. If the amount of harm done by an agent in violating the rationing restrictions is a function of how many other people violate them, the relevant description of his action might be "violating the restrictions when it is necessary to conserve gas for the war effort, and when the majority of people obey the restrictions." Whether or not reference to the actions of other agents should be included in an action's description has long been a point of contention in discussions of utilitarian generalization; in his criterion of relevance, Lyons provides us with a principled way of resolving this question in each particular case.

2. CRITIQUE OF LYONS' ARGUMENT FOR HIS CRITERION

Before we consider the import of Lyons' method of describing actions, it is important to note that the argument which purports to establish his criterion contains grave deficiencies – deficiencies serious enough to license

rejection of Lyons' apparent claim that his interpretation of UG is the only correct one.⁵

The argument in favor of Lyons' criterion may be stated more formally as follows:

Argument A

- Premise 1: Pure teleological principles are only concerned with the values of consequences of actions.
- Premise 2: In the context of such principles, acts may only be described in terms of their consequentially significant properties.
- Premise 3: UG is a pure teleological principle.
- Conclusion: The proper description of an action in the context of UG must include all and only the consequentially significant properties of that action.

Lyons believes that Premise 1 entails Premise 2.

The first difficulty with Argument A concerns the truth of Premises 1 and 3. The conjunction of these premises implies that UG is solely concerned with the values of consequences of actions. But is this true? Presumably, a principle is 'concerned' with certain properties of an action just in case its evaluation of the moral status of that action depends on its possession of those properties. Now, UG is applied by a two-stage process: first, the action in question is assigned to a set of relevantly similar actions, and second, the utility which would be produced by the performance of that set of actions is ascertained. Thus UG is concerned with *two* properties of any action – (1) the property in virtue of which it has membership in the set of relevantly similar actions, and (2) the property which expresses the value of the consequences that would be produced by the performance of that action (together with other actions in the set). Since we cannot assume these two properties to be identical, it is false that UG is *solely* concerned with the values of consequences of actions. But, since Premises 1 and 3 jointly imply that UG is strictly concerned with this aspect of any action, at least one of these premises must be false, and Argument A unsound.

Moreover, Argument A is invalid. As we have seen, Lyons claims that Premise 1 (Pure teleological principles are only concerned with the values of consequences of actions) entails Premise 2 (In the context of such

principles, acts may only be described in terms of their consequentially significant properties). This is incorrect. Premise 1 in fact implies that in the context of pure teleological principles, acts may only be viewed in terms of the *values* of their consequences – not in terms of the properties *in virtue of which* they produce those values. For evidently a pure teleological principle (if we accept Premise 1) is concerned only with the amount of utility an action produces, not with the manner in which it is produced. Thus, in terms of our previous example, it would be appropriate (in the context of a pure teleological principle) to describe the action of violating the gas-rationing restrictions as 'producing +10 utiles' (which is not a consequentially significant property, since it refers to, but does not cause, any valued effects), rather than as 'violating the restrictions when it is necessary to conserve gas...'. Hence even if we accept Premise 1, the inference from Premise 1 to Premise 2 is unwarranted, and Argument A invalid. The supposed fact that UG is solely concerned with the values of consequences of actions would not demonstrate Lyons' criterion of relevance to be the correct one, but rather would establish a different criterion requiring an action to be described in terms of the utility it produces.⁶

We have thus seen that Lyons' argument in favor of his criterion for describing actions contains at least one false premise and an invalid inference. The argument fails to demonstrate the correctness of Lyons' criterion, and thus provides no support for the proposed interpretation of UG which incorporates this criterion. Nevertheless, this version of UG is intrinsically attractive. Certainly it is one of the few adequately concrete interpretations available in the literature. Lyons' claim that this version is coextensive with AU therefore merits investigation.

3. NON-COEXTENSIONALITY BETWEEN UG AND AU

Lyons' argument that UG is extensionally equivalent to AU may be stated briefly as follows:

Argument B

- Premise 1: Relevantly similar actions produce equal utility.
- Premise 2: The generalized utility of an action is equivalent to the simple utility of the action multiplied by the number of relevantly similar actions.

Premise 3: For any set of relevantly described alternative actions, the value of n (the number of relevantly similar actions) is a constant.

Conclusion: UG and AU are extensionally equivalent.⁷

In Lyons' terminology, the 'generalized utility' of an action is the utility which would be produced by the universal performance of that action, while the 'simple utility' of an action is the utility of the consequences produced by that action alone. Lyons offers a convincing *reductio* argument (based on his criterion for describing actions) in support of Premise 1, claims that Premise 1 entails Premise 2 (p. 77), and then presents an independent argument in support of Premise 3. Although the unqualified claim that Premise 1 entails Premise 2 is false, as will be shown in a note,⁸ what I wish to do now is concentrate on the argument offered in support of Premise 3. Before looking at this argument, let us note (as Lyons does) why Premise 3 is necessary. Premises 1 and 2 are not sufficient to ensure that the value of n is a constant. Thus these premises by themselves are compatible with examples in which n varies. For example, suppose that an agent, Smith, has two alternative actions open to him, acts a and b . The performance of a would produce a utility of 5, while the performance of b would produce a utility of 4. Exactly three agents (including Smith) have the opportunity to perform acts relevantly similar to act a , while six agents have occasion to perform acts relevantly similar to act b . This case may be represented in the following Figure:

Alternatives	Simple Utility	Number of Similar Actions	=	Generalized Utility
a	5 (×)	3	=	15
b	4 (×)	6	=	24

Fig. 1. Case 1.

According to these figures, AU prescribes act a to Smith, while UG prescribes act b . But if cases like this can occur, then UG and AU are not coextensional, for they do not always yield the same prescriptions to the agent.

If n were a constant in this case, UG would prescribe a to Smith, and the prescriptions of the two theories would be identical. For example, if $n=3$ for both acts a and b , then the generalized utilities of the two actions would be 15 and 12 respectively, and UG, as well as AU, would prescribe act a . Thus, in order to establish coextensionality between AU and UG, the truth of Premise 3 must be demonstrated.

What argument can be offered in support of Premise 3? Lyons claims that his criterion for the correct description of actions *implies* that two actions are relevantly similar to each other only if they are members of similar alternative sets – i.e., that two actions a and a' are relevantly similar only if when a has an alternative b , then a' has a similar alternative b' . His argument for this claim is stated in the following passage:

... although the value of n can vary, it does not follow that in such cases, among a given set of alternatives, it is not a constant. I would argue that it is a constant, in effect, in such cases. Notice that each of the acts, a_1, a_2, \dots, a_m , has a complete general utilitarian description ' A_1 ', ' A_2 ', ..., ' A_m '. But it is an essential part of the circumstances surrounding a_1 , for example, that acts like a_2, \dots, a_m are open. The causally and teleologically significant features of the circumstances surrounding a_1 would be different if (some of) these alternatives were not open and others were. In other words, included (or implicit) in the description ' A_1 ' of a_1 is a reference to the *kinds* of alternatives open. The descriptions, ' A_1 ', ' A_2 ', ..., ' A_m ' are *internally related*. From this it follows that, for any given set of m alternatives, the value of n for the several kinds of acts is the same when the acts are completely relevantly described. (p. 117)

The contention here is that the availability of certain alternatives to an agent is causally relevant to the consequences which he actually brings about. Thus (on Lyons' criterion for describing actions), the relevant description of an action must refer to those alternatives. If the alternatives were different, then the description would necessarily be different also, and vice versa. It would follow from this that every act which bears a given relevant description has precisely the same (relevantly described) alternatives. If so, for any given set of actions, the number of similar actions (n) is a constant, as Premise 3 asserts.

Can we accept the claim that the relevant description of an action must include or imply reference to the range of alternatives open to the agent? Is it true that the causally significant features of the circumstances surrounding an action would be different if the alternatives to that action were different? Evidently, the best way to refute this claim is by to produce a counterexample.

Suppose that two actions, a_1 and a_2 , are performed respectively by Smith and Jones. Each action is a case of turning on a light by flipping a switch. The consequentially significant properties shared by these acts are properties such as occurring in the presence of a light switch, occurring in circumstances where the light switch is connected by wires to a live light bulb, occurring in circumstances where the agent will use the light to read a good detective story, and so forth. Let us imagine the two actions to have exactly similar results, e.g., the agent's experiencing 5 utiles while reading his book. Now let us introduce a new feature into the case: one of the circumstances present, when Smith performs his act a_1 , is the location of another switch next to his light switch. This other switch is connected to an exhaust fan. Since the fan switch is adjacent to the light switch, one of Smith's alternatives at the time he performs a_1 is action b_1 , turning on the fan. Let us further suppose that *no* such fan switch is located near Jones' light switch. Thus Jones does not possess the alternative of performing an action relevantly similar to b_1 . Despite their superficial similarity, acts a_1 and a_2 do not belong to similar alternative sets.

Must we conclude from this that acts a_1 and a_2 are *not* relevantly similar – that a_1 has some consequentially significant property that a_2 lacks? Lyons would have us believe this, but there seems to be no good reason for doing so. We can easily imagine that the 'causally and teleologically significant features of the circumstances' surrounding a_2 would *not* have been different, even if Jones' range of alternatives had been different – if, for example, he too had the opportunity of turning on a fan. It seems clear that (except in an extraordinary case) the utility of his action would be just what it is now, and the properties of his action which produced this utility would be exactly the same, even if he could have flipped a fan switch instead. Thus Lyons' claim that relevantly similar actions must belong to ranges of similar alternatives seems eminently implausible. Two actions, similar in all their consequentially significant properties, *can* have different sorts of alternatives. But it follows from this that n is not necessarily constant for any given range of alternatives, and hence that Premise 3 in Lyons' argument for the coextensionality of AU and UG is false. As we saw earlier, premise 3 is required to establish extensional equivalence. Since it is false, cases similar to Case 1 may occur, and UG is not coextensional with AU.⁹

This result should not cause rejoicing among advocates of UG, for Lyons' version, when properly understood, is clearly an unsatisfactory moral theory. It is easy to construct examples in which this version of UG generates extremely counter-intuitive prescriptions, precisely because it counts actions which are members of different ranges of alternatives as relevantly similar to the act in question. For example, consider a case in which an investigating officer must extract crucial information from a prisoner by fair means or foul. Since the prisoner refuses to divulge this information voluntarily, the officer must choose between torturing the prisoner or injecting him with truth serum. The only method of torture available to the officer is so peculiar that conditions permitting its application will never be duplicated. However, large numbers of persons will have the opportunity of injecting truth serum in prisoners with similar results. The utility figures for this case might then be as follows:

Alternatives	Simple Utility	Number of Similar Actions	Generalized Utility
Torture	-90(×)	1 =	- 90
Truth Serum	- 1(×)1000	=	-1000

Fig. 2. Case 2.

In this case, the application of UG (under Lyons' interpretation) yields the prescription to perform the torture. But this is outrageous – if the officer can obtain the information by either technique, he has a clear obligation to spare the prisoner as much pain as possible. The recommendation of UG cannot be accepted, and UG itself (under Lyons' interpretation) must be rejected as an unsatisfactory moral theory.¹⁰

We have shown that (a) contrary to Lyons' claim, UG (under his interpretation) is not extensionally equivalent to AU, and (b) this version of UG, although distinct from AU, does not constitute an intuitively acceptable moral theory. Both of these facts clearly derive from a single feature of Lyons' interpretation of UG – namely, its failure to guarantee the constancy of n throughout any given range of alternative actions. Once this defect is recognized, it is not difficult to formulate a version of UG from which it is eliminated.

4. A REVISED VERSION OF UTILITARIAN GENERALIZATION

Lyons' method of describing actions is easily amended to ensure that the description of any action must refer to the other alternatives available to the agent. Nor is this amendment completely foreign to the spirit of Lyons' approach. For, as we have seen, Lyons claims that UG is a 'pure teleological principle' and therefore solely concerned with the values of the consequences of actions. But if we accept this, it is clear that UG is a member of an important *subclass* of pure teleological principles – namely, the comparative principles. A comparative teleological principle assesses an action by comparing its (simple or generalized) utility with that of its alternatives. Thus such a principle is not simply concerned with the values of the consequences of actions, but rather with the *comparative* values of the consequences of actions. It might therefore be argued that, in the context of such principles, an action should be described, not simply in terms of its consequentially significant properties, but rather in terms of these properties together with a reference to its alternatives (also described in terms of such properties). Thus Lyons' general line of thinking should have led him to adopt the criterion for action description which I shall propose, rather than the one he actually opts for.

What we need are action descriptions which include (a) reference to the consequentially significant properties of the action, and (b) reference to the action's alternatives, similarly described in terms of their consequentially significant properties. This can be arranged most conveniently by introducing two 'levels' of action types (where an action type is understood to be a property that can hold true of individual actions).¹¹ Action types of the first level are defined in terms of the consequentially significant properties of the action, and action types of the second level are defined in terms of the level-1 type of the action, taken in conjunction with the level-1 types of its alternatives. Consider, for example, a case in which an agent possesses two alternatives, acts *a* and *b*. The definitions of the level-1 types of these actions would appear as follows:

Level-1 Definitions

act *a* is of type *A* at level 1 =_{df.} *a* has consequentially significant properties P_1, P_2, \dots, P_k .

act *b* is of type *B* at level 1 =_{df.} *b* has consequentially significant properties Q_1, Q_2, \dots, Q_k .

The definitions of the level-2 types of these actions would then appear as follows:

Level-2 Definitions:

act *a* is of type *A'* at level 2 =_{df.} *a* is of type *A* at level 1, and *a*'s sole alternative is of type *B* at level 1.
act *b* is of type *B'* at level 2 =_{df.} *b* is of type *B* at level 1, and *b*'s sole alternative is of type *A* at level 1.

If we now stipulate that the relevant description of an action specifies which level-2 action type is true of it, the description of the action will guarantee that it belongs to a determinate range of alternatives. For example, suppose that an agent performs an action of level-2 type *A'*. Then it must be the case that he had available exactly one alternative, namely an action of level-2 type *B'*. For the fact that he performed an act of level-2 type *A'* implies that he performed an act of level-1 type *A*, and had as his sole alternative an act of level-1 type *B*. But an act of level-1 type *B* in this context simply *is* an act of level-2 type *B'*. Similarly, if an agent performs an action of level-2 type *B'*, it must be the case that his sole alternative was an action of level-2 type *A'*. Thus acts of level-2 types *A'* and *B'* always 'come together', and the number of agents who have the opportunity of performing an act of level-2 type *A'* is identical with the number of agents who have the opportunity of performing an act of level-2 type *B'*. By adopting this criterion for action description, we have ensured that the number of relevantly similar actions in any alternative set is a constant.

Using this revised criterion for the description of actions, we eliminate at one stroke both the source of non-coextensionality between AU and the previous version of UG, and also the source of the counter-intuitive prescriptions generated by that version of UG. May we conclude that this new version of UG is coextensional with AU, or, if not, that it is a sound moral principle, worthy of adoption in AU's place? Unfortunately nei-

ther of these conclusions is warranted, as we shall discover in the next two sections.

5. UTILITARIAN GENERALIZATION IN CHOOSING FUTURE ACTIONS

Is UG (under the new interpretation) a sound moral principle? I shall argue that it is not. As background to this argument, let us notice that any moral theory may serve either (or both) of two distinct functions. First, an agent may use a theory in deciding which action he should perform. Thus, a theory may generate prescriptions for *future* actions. Second, an observer may use a theory to determine whether or not the action actually performed was the right action. That is, a theory may be a criterion of the rightness and wrongness of actions.¹² Normally we do not distinguish these two functions, but on occasion it is necessary to do so (for example, some theories may only be employed in one of these capacities). In this section I shall argue that UG is incapable of performing the first of these functions, and in the next section claim that it also cannot perform the second, for closely related reasons.

In order to see how UG may not be used by an agent in choosing which action to perform, we must apply UG to a particular case. I shall describe a simple voting example, chosen because it requires the description of an action to refer to the activities of other agents, and also because UG is frequently defended in association with more complex cases having the same structure.

Consider three agents, Smith, Jones, and Brown, who are neighbors in an apartment house and share a yard. They agree to split the expenses of any play equipment to be installed in the yard for the use of their children. According to their agreement, the majority rules in such decisions: if two of them wish to purchase a certain piece of equipment, the other is bound to pay his share of the expenses. In the case at issue, question has been raised whether or not a swing set should be purchased for the children. If it is, the utility produced will be 10; if not, the utility produced will be 3. The agents vote simultaneously. Figures for the case appear in the following matrix: (See p. 89).

Consider Rows 7 and 8 of this matrix. These rows represent a situation in which both Smith and Jones vote 'no' and Brown must choose between voting 'yes' and voting 'no'. In order that Brown's decision is not

	Smith	Jones	Brown	Utility of Total Consequences
1.	yes	yes	yes	10
2.	yes	yes	no	10
3.	yes	no	yes	10
4.	yes	no	no	3
5.	no	yes	yes	10
6.	no	yes	no	3
7.	no	no	yes	3
8.	no	no	no	3

Fig. 3. Case 3.

made in ignorance, let us add that he knows that both Smith and Jones will vote 'no'. Now, in order to apply UG to the case and derive a recommendation for Brown, we must describe his alternatives in terms of their level-2 act types, types *N* and *Y*. Since the consequences of Brown's actions depend on how the others vote, his act of voting 'yes' has the following consequentially significant properties (perhaps among others): it is a 'yes' vote, it occurs in a context where two 'no' votes are cast, and it occurs in a context in which the casting of two 'no' votes is sufficient to ensure the occurrence of events having a total utility of 3. The definition of level-2 act type *Y* must include reference to all these properties. Similarly, his act of voting 'no' has the following consequentially significant properties (perhaps among others): it is a 'no' vote, it occurs in a context in which two other 'no' votes are cast, and it occurs in a context in which the casting of two 'no' votes is sufficient to ensure the occurrence of events having a total utility of 3. The definition of level-2 act type *N* must include reference to these properties.

Suppose now that Brown attempts to use UG in deciding what he should do. In order for Brown to ascertain which action UG advises him to take, he must first identify the set of agents who have the opportunity to perform relevantly similar acts (i.e., acts of types *Y* and *N*). Having identified these agents, he can then calculate the consequences of everyone's performing these alternatives, and on this basis derive a recommendation from UG on which action to perform.

Which agents do have the opportunity to perform acts of these types? In particular, which agents have the opportunity to perform an act of type *Y* (roughly, which agents are in a position to vote 'yes' when two 'no' votes are cast)? Until Brown himself acts, *there is no determinate answer to this question*. If Brown should vote 'yes', then neither Smith nor Jones will have had the opportunity to vote 'yes' when two 'no' votes are cast. Only Brown himself will have had this opportunity, so the set of agents possessing this option only includes one member. But if Brown should vote 'no' instead, then *both* Smith and Jones (as well as Brown) will have had the opportunity of voting 'yes' when two 'no' votes are cast. Thus, in these circumstances, the set of agents possessing this option includes three members. Hence, before Brown himself acts, there is no determinate set of agents who have the opportunity of performing acts of type *Y*, for Brown's eventual action forms part of the circumstances which determine the act types of the other agents' actions. Since there is no determinate set of agents who have this opportunity, there is no way to ascertain what the consequences would be if everyone performed an action of type *Y*. Since he cannot determine the generalized utility of one of his alternatives, there is no way for Brown to employ UG in choosing which action he should perform.¹³

This example is only one of a wide range of similar cases – precisely the cases, in fact, to which UG has standardly been applied by its advocates with great apparent success: cases in which an agent must decide whether or not to vote for his candidate, or to violate gas-rationing restrictions during wartime, or to cut across the park lawn. Each of these cases involves the production of what Lyons terms 'threshold effects' – consequences that are only produced when a certain critical density of similar actions is attained. The election of a candidate is a threshold effect, for example, because it only occurs when 51% of the votes are cast in his favor. Thus, in these standard cases, the consequences of one agent's action heavily depend on the activities of other agents – how many others vote, or obey the rationing restrictions, or walk around the lawn. Hence the activities of others must be referred to in the description of each agent's action. Patterns of action may occur like the one just described, such that the agent is prevented from employing UG to decide which action he should perform. Thus, contrary to the hopes of its proponents, UG may not be of any help to an agent in deciding whether to

vote, or to use more than his share of gas, or to walk across the grass. We have not simply discovered one isolated counterexample to the use of UG in prescribing future action, but rather an entire range of counterexamples which are made all the more significant by the fact that they are the very cases which UG was originally introduced to resolve.

6. UTILITARIAN GENERALIZATION AS A CRITERION OF RIGHTNESS AND WRONGNESS

We have now seen that UG (under the revised interpretation) cannot be used to prescribe future action in a number of important cases. We will now discover that it cannot be used as a criterion of rightness and wrongness, for there are cases in which it fails to determine whether or not an action is right, even after the action is performed, and all the facts are in.

Let us continue to examine the example just described. Suppose the actual situation is that represented by Row 8 of the matrix – all three agents vote 'no'. In these circumstances, the alternatives of all three agents are relevantly similar: each performs an action of type *N*, and, given the fact that they all vote 'no', each has the alternative of performing an act of type *Y*. To ascertain whether or not Brown does right in voting 'no', we must compare the generalized utility of his voting 'no' with the generalized utility of his voting 'yes'. If we assume that these three agents are the only ones who will ever have occasion to perform acts of type *N*, the generalized utility of Brown's voting 'no' is 3 (the utility actually produced in this situation). But what is the generalized utility of his voting 'yes'? This is the utility which would be produced by the performance by all three agents of acts of type *Y*. But an act of type *Y*, as we saw before, is an act partially defined as "voting 'yes' when two 'no' votes are cast." And, despite the fact that each of these three agents has the alternative of performing this sort of act, it is impossible for all three *together* to perform these acts! For performance of any such act requires the agent to vote 'yes', but if all three agents vote 'yes', then it is not true of any one that he is voting 'yes' when two 'no' votes are cast. Given the nature of the actual circumstances, it is not possible for all three agents to perform their alternatives of type *Y*.

If it is impossible for all three agents to perform this action, then there is no determinate set of consequences that would result if they were to do

so, and thus no determinate utility to the universal performance of acts of type *Y*. In these circumstances, the generalized utility of Brown's act of voting 'no' cannot be compared with the generalized utility of his alternative act, and UG does not specify which of his alternatives is right.¹⁴

Our earlier case therefore provides us with an example in which UG cannot function as a criterion of rightness and wrongness, since it ascribes neither rightness nor wrongness to this agent's action. As we pointed out before, the example is simply a less complex version of the sort of case to which UG has traditionally been applied by its advocates. Similar patterns of action may occur in these standard cases, in which the generalized utility of an agent's alternative would necessarily be indeterminate. UG would then provide no answer to the question of whether or not it is right of the agent to vote for his candidate, or to violate the gas rationing restrictions, or to walk across the park grass. Our argument shows that UG (under the revised interpretation) cannot serve as a criterion of rightness and wrongness in a broad range of cases, the very cases which advocates of UG have hoped to show could only be handled satisfactorily by their theory.¹⁵

Having seen that UG generates no prescription in our voting example, it is easy to show that this version of UG is not coextensive with AU. Unlike UG, AU prescribes an action to Brown, namely the action of voting 'no', since this action has a simple utility of 1 as compared with the simple utility of 0 produced by his alternative of voting 'yes'.¹⁶ Since AU generates a prescription in this case and UG generates none, they are not extensionally equivalent.

7. CONCLUSION

We have now examined two versions of UG – that proposed by David Lyons, and a new version derived from his. Both were found to be extensionally nonequivalent to AU, but both were also found to be unacceptable moral theories (Lyons' version generated counter-intuitive prescriptions, and the derived version failed to generate any prescriptions at all in an important range of cases). Naturally, these results cannot be taken to show that no satisfactory form of UG will be forthcoming. But at least some of the conditions to be met by future proposals are now

clear: any successful version of the theory must interpret the 'generalized performance' of an action to include only the actions of agents who face similar ranges of alternatives, and (in some cases) must select as relevant a description of the action which does not refer, either explicitly or implicitly, to the actions of other agents. Reflection shows that the same remarks hold true, *mutatis mutandis*, for rule utilitarianism as well: unless the rules are formulated with extreme care in these respects, they may yield unacceptable recommendations, or else fail to recommend any action at all in the very cases they are designed to resolve.¹⁷

University of Pittsburgh

NOTES

¹ David Lyons, *Forms and Limits of Utilitarianism*, Oxford University Press, London, 1965. References to this work will henceforth appear in parentheses within the text.

I would like to thank Richard Brandt, William Frankena, Alvin Goldman, and W. D. Hart for their comments on earlier versions of this paper. Most of this material originally appeared in my doctoral dissertation, *The Generalization Principle in Ethics*, University of Michigan, August, 1972.

² There is another reason (which Lyons does not consider) for preferring utilitarian generalization to act utilitarianism. That is, one might require an adequate moral theory not only to generate prescriptions which match our intuitions about particular cases, but also to reflect the structure of the reasoning we engage in while arriving at those intuitions. Thus, utilitarian generalization might be superior to act utilitarianism, even though they are coextensional, if the former theory corresponds more closely to the actual principles we employ in drawing conclusions about what it would be right to do.

³ This formulation differs from the one offered by Lyons, which is stated as follows: "If and only if the consequences of everyone's doing a certain sort of thing would be worse than those of some alternative, then it would be wrong for anyone to do such a thing." (p. 25) Although the difference is negligible, my version makes it clear that we are to compare the consequences of everyone's doing a certain sort of thing with the consequences of everyone's performing any alternative sort of action (rather than comparing the former consequences with the consequences produced by *any* alternative pattern of action, including non-homogeneous patterns).

⁴ It is equally important to assign some interpretation to the term 'consequences' in both principles. Although Lyons does not explicitly note this fact, he appears to adopt the suggestion that if an action *a* together with *k* actions of other agents causes an event *e*, then $1/k + 1$ th of *e* is a consequence of *a*. (p. 85) If we interpret AU in terms of this notion of consequences, it is easy to construct cases in which the action prescribed by AU is followed by events having less utility than the events which would follow some other action available to the agent. Since this is objectionable, it would be better to adopt a notion of consequences according to which an event is a consequence of an action just in case that event would occur subsequent to the action, but would not occur subsequent to the performance of at least one of the action's alternatives. However, the arguments presented in the body of this paper do not depend on the

precise notion of consequences used (although some examples would require small changes to accommodate the proper notion). To preserve contact with Lyons' position, I retain his interpretation of this term.

⁵ I say this is an 'apparent' claim because later, in Chapter V (p. 161), he seems to reject it.

⁶ Whether or not UG (when interpreted with this new criterion) would yield different prescriptions depends on issues taken up in the next section. In any event, this is clearly not the criterion envisioned by early proponents of UG. See note 10.

⁷ Chapter III, *passim*.

⁸ Premise 1 appears to imply Premise 2 when taken in conjunction with the notion of consequences adopted by Lyons (see note 4). Premise 2, however, is false when the proper notion of consequences is used. For example, consider a case in which 10 people all perform relevantly similar acts, each of which is necessary for the joint production of an event having a utility of 100. On Lyons' notion of consequences, the simple utility of each action = 10 [$1/10 (\times) 100$], and the generalized utility of each action is 100 [$10 (\times) 10$], as Premise 2 requires. But on the proper notion of consequences both the simple and the generalized utilities of any one action = 100, contrary to Premise 2. However, since some linear equation describes the relation between the action's simple and generalized utilities, this fact need not disturb us, since that is all Lyons' proof really demands.

⁹ After developing this point, I have discovered that the same conclusion has been reached by Gregory Kavka in his unpublished paper 'Extensional Equivalence and Utilitarian Generalization' and by Fred Feldman in his forthcoming paper 'On the Extensional Equivalence of Simple and General Utilitarianism'.

¹⁰ We can now see that the new criterion for describing actions suggested in Section 2 would not generate a version of UG which is either coextensional with Lyons' version, or an adequate moral theory, since it, too, does not require an action's alternatives to be mentioned in the relevant description of that action. However, a related criterion which stipulates that actions be described in terms of their comparative utilities would circumvent this problem. Unhappily, a version of UG employing this criterion is all too obviously coextensional with AU.

¹¹ This technique was suggested to me by W. D. Hart.

¹² The theoretical importance of making this distinction was called to my attention by R. Eugene Bales, 'Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?', *American Philosophical Quarterly* VIII (1971), 256-65.

¹³ Of course, if determinism is true, then there is a sense in which, before Brown acts, there is a determinate set of agents who have the opportunity to perform an act of type Y. However, since Brown is making a *decision* about what to do, we cannot assume that *he* knows before this decision is made which act he will perform.

¹⁴ After developing this point, I have discovered that substantially similar conclusions have been reached by Allan F. Gibbard in his unpublished doctoral dissertation, *Utilitarianism and Coordination*, Harvard University; Howard Sobel, 'Utilitarianisms: Simple and General', *Inquiry* 13, 420-21; and by Gregory Kavka in his paper 'Extensional Equivalence and Utilitarian Generalization'.

¹⁵ In 'A Defense of Rule-Utilitarianism Against David Lyons', *Journal of Philosophy* LXV (1968), Gertrude Ezorsky has pointed out a related difficulty with Lyons' version of UG. She proposes to avoid this problem by eliminating reference to the activities of other agents in the descriptions of problematic actions. Her first suggestion is simply to drop such references from the complete Lyons-description of the action. Thus, in effect, we would describe Brown's action as "voting 'yes'." Obviously, all three of our

agents can perform this action. However, Ezorsky points out that such descriptions lead to counter-intuitive prescriptions in cases where minimizing-conditions obtain (*i.e.*, where so few people cooperate that the agent cannot bring about the threshold effect whatever he does), and advances the alternative suggestion that we describe actions like Brown's as "voting 'yes' when a 'yes' vote will not bring about the purchase of the swing set." But it is clear that this description of the act has the same problem that our original one does; for even though all three agents have the opportunity to perform an act of this type, it is not possible for all to actually perform it (for if they all vote 'yes', then the swing set *will* be purchased).

¹⁶ These utility figures are derived on the basis of Lyons' notion of consequences. On the proper notion of consequences, AU prescribes either action as right.

¹⁷ Disillusioned proponents of rule utilitarianism might note that one of Lyons' central arguments concerning *this* theory is also mistaken. In Chapter IV, Lyons considers a version of RU entitled 'Specious Rule Utilitarianism' which is formulated as follows: "An act is right if, and only if, it conforms to a set of rules *general conformity* to which would maximize utility." Lyons claims that SRU is incapable of discriminating between two sets of rules identical except that one includes minimizing-conditions and the other does not. If this is true, then both sets of rules, according to SRU, generate correct prescriptions (assuming their 'conformity-utilities' are better than those of any other set of rules). But, since they prescribe different actions in a situation where minimizing conditions actually obtain, SRU generates incompatible prescriptions. Lyons' famous argument for this thesis is stated as follows: "... the general conformity test will not distinguish between these two sets. For the test is based upon the supposition of general conformity. Under such a supposition, the minimizing-conditions will simply be vacuous. For the minimizing-conditions are satisfied only when there is not general conformity. Thus no difference will show up between the rules that contain minimizing-conditions and the rules that do not, under the general conformity test. This means that (SRU) is indeterminate for this class of cases." (p. 138)

What Lyons assumes in this argument is that two rules *R* and *R'*, identical except for the inclusion of minimizing conditions in *R'*, have equal conformity-utilities, because the pattern of activity which qualifies as general conformity to *R'* is the *very same* pattern of activity which qualifies as general conformity to *R*. If the patterns of activity are the same, then naturally the utilities produced will be identical as well. Lyons believes that the patterns of activity will be the same because he believes that whenever *R'* is generally conformed to, its 'minimizing-conditions are simply vacuous'. However, this is simply false. Consider, for example, the rules *R* ('Always vote for your candidate') and *R'* ('Always vote for your candidate unless he is going to lose no matter whether you vote or not'). *R* is generally conformed to in a case where everyone votes for his candidate. Similarly, *R'* is generally conformed to in a case where everyone votes for his candidate. But *R'* is also conformed to in a case where *no one* who favors a certain candidate votes for him at all (for then it is true of each such person that his candidate will lose no matter what he does, and hence each one is following *R'*).

Thus the utility produced by general conformity to *R* is not necessarily equivalent to the utility produced by general conformity to *R'*. However, SRU is still indeterminate, for there is no determinate utility produced by general conformity to *R'* (or any other rule containing minimizing conditions), since there are several *different* ways in which *R'* can be generally conformed to. Thus there is no coherent way to compare *R'* to other rules, and SRU selects no rules at all, since the required comparisons among all candidates cannot be made.