

BOSTON UNIVERSITY SCHOOL OF LAW

WORKING PAPER SERIES, PUBLIC LAW & LEGAL THEORY
WORKING PAPER No. 99-7



THE MORAL OPACITY OF UTILITARIANISM

DAVID LYONS

THIS PAPER WILL APPEAR IN RULE CONSEQUENTIALISM: A READER, ED. BRAD HOOKER, ELINOR MASON & DALE E. MILLER (FORTHCOMING FROM THE **EDINBURGH UNIVERSITY PRESS**)

This paper can be downloaded without charge at:

The Boston University School of Law Working Paper Series Index:
<http://www.bu.edu/law/faculty/papers>

The Social Science Research Network Electronic Paper Collection:
http://papers.ssrn.com/paper.taf?abstract_id=212288

THE MORAL OPACITY OF UTILITARIANISM¹

David Lyons

I. The Problem

Utilitarians sometimes suggest that their moral theory has an advantage over competing theories in basing moral judgments on the consequences of conduct. As its dictates are determined by empirically determinable facts, it offers a procedure for settling moral controversies on objective grounds. One need not appeal, for example, to the dubious authority of “moral intuitions.”²

Claims like these are subject to familiar objections at various levels. I shall mention a representative sample and then focus on more serious difficulties stemming from aspects of utilitarianism that I believe have not been fully enough explored.

The claim made on behalf of utilitarianism seems premature. First, the suggested contrast between utilitarianism and its competitors obscures a more fundamental similarity. For we have no reason to believe that a utilitarian conception of right conduct could be established as sound solely by reference to empirically confirmable facts. Second, suppose we assume a consequentialist framework. The question, what consequences have value, does not present itself as an ordinary empirical issue. Thirdly, suppose we take for granted that utilitarianism concerns itself specifically with the welfare or well-being of individuals. This is a standard reading of utilitarianism and by no means an ungenerous interpretation. Questions about the nature or constituents of individual welfare do not seem decidable by ordinary empirical means.³

Each of the considerations so far mentioned suggests that value judgments are required for determining the utilitarian criterion of right conduct. Until the criterion is identified, we do not have a theory that is available for application.

I turn now to complications of a different kind: they assume there are solutions to the foregoing problems and initially concern the application of a utilitarian standard.

Determining the utilitarian difference it makes to act one way rather than any other is at best extremely difficult. No more than one of the alternative courses of conduct that are

¹Ancestors of this paper were presented at Boston University, Carleton University (Ottawa), SUNY Albany, SUNY Geneseo, the University of Colorado, the University of Texas, and the International Society for Utilitarian Studies. I am grateful to discussants on those occasions, and especially to Alan Fuchs, Thaddeus Metz, William Nelson, Martino Traxler, and Henry West.

²The problems which I shall later discuss concern both criteria of right conduct and practical procedures for deciding what conduct is morally permissible, so I shall generally ignore that important distinction.

³Nor does the question, how to take account of the welfare of nonhuman animals, seem an ordinary empirical one either.

available in a given situation can be initiated, for example, and some of its potentially relevant consequences will always remain unrealized. In practice, furthermore, judgments of marginal utility are subject to bias and wishful thinking. So a utilitarian principle may not provide a very reliable decision procedure.

I want now to suggest a closely related but much more serious set of problems for utilitarianism. The discussion so far assumes, in effect, that utilitarianism either is or in principle grounds a single, determinate moral theory -- one criterion of right conduct. Moral theorists already recognize that, in one sense at least, this is not so. For it is commonplace to distinguish act and rule utilitarianisms and more generally direct and indirect forms of the theory. We have not yet, I think, fully appreciated how deep the potential ambiguity may go.

In this paper I shall explain, first, why the moral implications of utilitarianism are at best **opaque** -- practically impossible to determine -- because the utilitarian criterion of right conduct is itself practically impossible to determine. That sort of indeterminacy can obtain even if, in the final analysis, utilitarianism determines one and only one criterion of right conduct. I shall suggest, secondly, that utilitarianism does not determine one unique criterion of right conduct; that is, no single criterion of right conduct best represents utilitarianism. If that is in fact the case, then utilitarianism is not just morally opaque but is morally **ambiguous** -- lacking determinate implications for conduct.

II. Act Utilitarianism

For much of the twentieth century, utilitarianism has been understood as preferring the approach to morality that requires individuals to act so as to promote welfare as much as possible; that is, to maximize utility. In other words, the utilitarian approach to morality has been equated with (some form of) act utilitarianism. This understanding may be less common today than it was a couple of generations ago, but I think it remains commonplace.

Act utilitarianism applies directly to particular acts as performed by particular individuals on particular occasions. It directly evaluates conduct, and only conduct, from a moral point of view. It does not directly evaluate other things that are also subjected to moral appraisal, such as laws, character traits, or moral attitudes, even though these are usually assumed to have utilities. Act utilitarianism takes account of those other utilities indirectly, by considering the effects of conduct on laws, traits, and attitudes and the effects of laws, traits, and attitudes on conduct.

The act utilitarian appraisal of acts does not seem to entail or to be entailed by utilitarian moral appraisals of other things, such as laws. Because acts that affect what laws we have possess other consequences as well, we cannot assume that optimific acts bring about optimific laws or that acts which bring about optimific laws are themselves optimific. This suggests that evaluations of conduct that primarily focus on acts or laws, respectively, do not necessarily converge. I'll return to this point presently.

In the recent past, other kinds of utilitarian moral principles have been developed, the most familiar being rule utilitarianism. The latter requires of individual acts, not that they maximize welfare, but that they conform to rules that maximize welfare.

Rule utilitarianism was developed, at least in part, so as to accommodate moral objections to act utilitarianism. It may therefore be regarded as a compromise between utilitarian and

nonutilitarian ideas about morality. Partly for that reason, act utilitarianism is still seen by many theorists as the most faithful moral expression of what Peter Railton calls “the guiding utilitarian idea ... that the final ground of moral assessment ... must lie in effects on people's well-being.” (Railton 1988, p. 398)

I will return later to the guiding utilitarian idea and the distinction between it and utilitarian moral principles. For the sake of argument now, I assume that the guiding utilitarian idea calls for welfare to be promoted as much as possible. For, if the realization of an increment of welfare is assumed to have positive value and to provide a reason for any course of conduct that brings it about, it would seem that the very same theory favors both maximizing welfare and any course of conduct that does so.

It is natural to suppose that act utilitarianism is the most faithful moral expression of the guiding utilitarian idea, so understood. For act utilitarianism requires one to promote welfare as much as possible. In other words, it is natural to suppose that perfect conformity with a principle that requires one to maximize welfare would, on any and all occasions, maximize welfare.

But we have reason to doubt that assumption. We have reason to believe that conforming our behavior to act utilitarianism does not guarantee that we shall maximize welfare. If that seemingly paradoxical claim is true, the question arises whether there is a better interpretation of utilitarianism -- a better representative within moral theory of the guiding utilitarian idea.

I shall now review briefly some reasons for thinking that act utilitarianism may not be the best moral interpretation of the guiding utilitarian idea. Then I shall explore the possibility of an alternative.

To simplify matters, I will consider only theories which require that welfare actually be maximized (actual-consequence utilitarianism). Perfect compliance with a probable-consequence version of act utilitarianism, which requires conduct that is most likely to maximize welfare, can fail to maximize welfare. This is possible because a justified estimate of consequences can be mistaken. No such complication prevents conformity to actual-consequence act utilitarianism from maximizing welfare. As the very formulation of act utilitarianism seems to guarantee, it is natural to suppose that welfare is promoted as much as possible on any occasion when actual-consequence act utilitarianism is followed perfectly.

If that assumption is mistaken, it would seem to mark a **utilitarian** shortcoming of act utilitarianism. The assumption seems mistaken.

III. Accessible but Unrealized Increments of Value

The first problem is this. The acts of two or more people might each conform perfectly to act utilitarianism without promoting welfare as much as it would be possible for them collectively to do, in the same surrounding circumstances, by means of a different pattern of conduct.

The following example suggests why. Suppose that Alice and Barbara sing together and improvise as they perform. They harmonize well, so it is better when they harmonize than when they merely sing together. The musical results are good when Barbara sings high and Alice sings low, but they are best when Alice sings high and Barbara sings low. If Alice sings high, then

(other things equal⁴) act utilitarianism requires Barbara to sing low, and if Barbara sings low, then it requires Alice to sing high. In that case, all of the (directly relevant) benefits that they might collectively realize through their singing will be realized. Suppose, however, that Barbara sings high: then act utilitarianism requires Alice to sing low; and if Alice should sing low, then it requires Barbara to sing high. In the latter two cases, some benefits which Alice and Barbara **might** realize through their conduct will **not** be realized.⁵

How much welfare is promoted depends not just on what Alice and Barbara do but also on the relations between their acts. Different patterns of conduct can result in different degrees to which welfare is promoted. Alice and Barbara might each maximize welfare and conform perfectly to act utilitarianism, given what the other does. But, in some cases, Alice and Barbara might collectively promote welfare to a greater degree if they both acted differently. Because of this, some available increments of welfare may **not** be realized by conduct that conforms **perfectly** to act utilitarianism.

So, while perfect conformity to act utilitarianism guarantees that the conduct of one or more individuals promotes welfare as much as it can in their respective circumstances, perfect conformity to that principle does not guarantee that welfare will be promoted as much as it might be by all of those individuals in their joint circumstances. This should lead us to wonder whether perfect conformity to a different principle would promote welfare more effectively.⁶

IV. Available but Inaccessible Value

Here's another problem. Some available increments of welfare may not be accessible at all to conduct that conforms perfectly to act utilitarianism. This problem derives from the fact that utility can be assigned to many things besides acts.⁷

Appreciation of this fact has given rise to novel variants of utilitarianism. We have

⁴This qualification will be understood but not stated hereafter.

⁵In Allan Gibbard's original example, the best outcome results from both persons' doing X, the next best from their both doing Y, and their acting differently is neutral. (Gibbard 1965) My variation on Gibbard's example makes the point that the relevant coordination patterns need not involve everyone's doing the same thing.

⁶Don Regan's "cooperative utilitarianism" is designed to address this problem. (Regan 1980) Regan's principle requires an individual to coordinate with others optimally whenever possible. It is unclear whether perfect conformity to that principle would in fact promote welfare better than would perfect conformity to act utilitarianism. We may never have adequate grounds for a confident answer to this complex question. In any case, cooperative utilitarianism does not attempt to address the next problem.

⁷This point was appreciated by Jeremy Bentham and John Stuart Mill. Both writers judged various objects of moral appraisal, aside from acts, such as laws, character traits, and moral attitudes, directly by reference to their respective utilities. From a twentieth century perspective, this might seem to show that, in their hands, utilitarian moral theory had not yet matured. That judgment need not be disrespectful. It would not disparage either writer to suggest that later theorists were enabled by their predecessors' work to develop a more refined moral interpretation of the guiding utilitarian idea.

We should withhold judgment, however: Bentham and Mill might have understood utilitarianism better than their actual utilitarian successors.

already noted the development of rule utilitarianism. Another variant is motive utilitarianism, which concerns behavioral dispositions such as character traits and subjective interests. As Robert Adams defines it, motive utilitarianism holds “that one pattern of motivation is morally better than another to the extent that the former has more utility than the latter.” (Adams 1976, p. 470)

Adams explains motive utilitarianism by using an example with the following structure. Jack has interests, pursuit of which realizes benefits that presuppose those interests. His interests lead him to some acts that fail to maximize value in their circumstances and accordingly violate act-utilitarianism. The important claim is this: If Jack lacked those interests, his conduct could conform to act-utilitarianism, but he would also realize a lower level of benefits in the long run than he would by pursuing those interests.

Railton observes that “certain goods are readily attainable -- or attainable at all -- only if people have well-developed characters.” (Railton 1984, p. 158) He argues that “loving relationships, friendships, group loyalties, and spontaneous actions are among the most important contributors to whatever it is that makes life worthwhile.” (p. 139) These factors may be so useful overall that a “sophisticated” utilitarian would generally favor their possession and nurturance, even though they may result in violations of act utilitarianism. (p. 153)⁸

The relevant point is that very useful dispositions may generate some acts that do **not** promote the general welfare as much as it is possible for the individual to do in the circumstances; but it may nevertheless be **more** useful in the long run to be led by those dispositions sometimes to violate act utilitarian dictates than to lack the dispositions and conform perfectly to act utilitarianism.

Here's an example. Honesty is a useful trait of character -- so useful, indeed, that in the long run welfare may be better served when people are honest, act in character, and as a consequence do not always act so as to maximize welfare, than it would be served if people conformed perfectly to act utilitarianism. If that is so, it is presumably because consistent manifestations of honesty in conduct encourage and facilitate cooperative projects that promote welfare. In short, firm honesty of character, manifested in conduct that is sometimes sub-optimal, can have benefits that exceed its costs.

The point I take from these examples is that some available increments of welfare may not be accessible to act utilitarian conduct. Perfect conformity to act utilitarianism may be causally incompatible with the realization of those benefits. Most important, welfare may be promoted to a greater degree if act utilitarianism is violated and those increments of welfare are realized.

For this possibility to be realized, another condition must be satisfied. It must not be possible for the relevant behavioral dispositions to be brought about by welfare-maximizing conduct. If that were possible, then the relevant increments of welfare might be accessible to act

⁸I have paraphrased part of Railton's argument so that it refers to utilitarianism. He initially makes his points relative to a broader consequentialism. (Railton 1984, pp. 149ff) The difference between consequentialism and utilitarianism does not matter for present purposes.

utilitarian conduct indirectly.⁹

The implications of the Adams and Railton examples, therefore, should not be exaggerated. They establish only the **possibility** that welfare is better promoted when conduct reflects useful behavioral dispositions that lead to violations of act utilitarianism than if it conforms perfectly to act utilitarianism. The examples do not show that all the relevant factual conditions are satisfied in the real world.

So we cannot infer from the examples that act utilitarianism is not the most faithful moral expression of the guiding utilitarian idea. But we cannot reasonably assume the contrary, either.

I am going to suppose that the examples are not misleading, and that act utilitarianism **may not be** the most faithful moral expression of the guiding utilitarian idea. That leaves us with the question, **What sort of utilitarian principle might best express the guiding utilitarian idea?**

V. Generic Utilitarian Reasoning

It is time to draw attention to some features of my argument so far. They concern a little noticed aspect of theorizing about utilitarian moral principles, which I shall call **generic utilitarian reasoning**.

Consider the familiar claim that using act utilitarianism as a practical guide to the morality of conduct is counter-productive because estimates of consequences are unreliable and more accurate guides to optimific conduct are available.

That argument can be understood in two different ways. It might mean that there are more dependable methods of complying with the requirements of act utilitarianism than by consciously and deliberately trying to follow that principle. On this reading, the claim is that using act utilitarianism as a decision procedure is counter-productive from an act utilitarian point of view.

But the argument is usually understood differently. Then it means that there are **better methods of promoting welfare** than by consciously and deliberately trying to follow act utilitarianism. Using act utilitarianism as a decision procedure is counter-productive, not from a specifically act utilitarian perspective, but from a generic utilitarian point of view.

The theoretical literature contains other examples of generic utilitarian reasoning about utilitarian moral principles. Consider a standard version of rule utilitarianism. It requires that we follow rules, general conformity to which would maximize utility. It is sometimes said, however, that one person's compliance with that rule utilitarian principle can be counter-productive because following some rules is useful only when many are following them. A standard example is resistance to an oppressive regime. Isolated resistance may have worse consequences than no resistance at all. In many circumstances, to be effective, resistance must be widespread.

Thus, small-scale compliance with utilitarian moral principles can be harmful rather than beneficial overall. This is a **utilitarian** argument against the rule utilitarian principle itself. It is

⁹One reason to doubt this is that the inculcation of behavioral dispositions requires extended patterns of conduct, the utility of early stages of which depends on continuation of such training.

another example of generic utilitarian reasoning about a utilitarian moral principle. It concerns whether or not conformity to the principle would promote welfare.

I shall not evaluate the factual claims made in generic utilitarian arguments. What is important for present purposes is not their soundness but that they assume a distinction between the guiding utilitarian idea and utilitarian moral principles. The guiding idea is not formulated in moral terms; nor does it entail any particular moral standard. It assumes that individuals' welfare has basic value and implies that more welfare is better than less. It holds that moral worth is relative to the promotion of welfare. But no premise of that sort entails a specific moral principle, such as one claiming that an act is morally wrong unless it maximizes welfare.¹⁰ The differences between the guiding utilitarian idea and utilitarian moral principles help to explain how a number of different principles can be counted as utilitarian.

I have assumed here that any standard using the criterion of maximizing welfare as the basis for moral evaluation counts as a utilitarian moral principle. The resulting class includes principles that appraise acts (more or less directly) by reference to their utilities, principles that appraise character traits by reference to their utilities, principles that appraise laws by reference to their utilities, and so on.¹¹

I have assumed, further, that a commitment to the value of welfare implies a commitment to maximizing utility. So, in asking what moral principle most faithfully represents the guiding utilitarian idea, I have understood this as the question, **“Can we identify a moral principle, conformity to which would most effectively promote welfare?”**¹²

VI. Valoric Utilitarianism

In grappling with the question I just posed, I have consulted others' attempts to address related problems. Peter Railton, for example, imagines the construction of a complex theory which he dubs “valoric utilitarianism.” (Railton 1988, pp. 409-413) The first step in constructing it is to apply the criterion of maximizing utility to all objects of moral appraisal. Character traits, persons, acts, rules, institutions, and resource allocations, for example, would each be ranked, on **generic** utilitarian grounds, as “more or less morally fortunate.” In the relevant version of this theory, such utilitarian rankings would not be confused with ordinary moral appraisals; for such an account of moral rightness would be implausible: there is a conceptual mismatch between scalar utilitarian rankings and binary moral appraisals like right or wrong. To achieve a valoric utilitarian account of right action, another step is needed:

once the valoric utilitarian moves beyond judgments of what is more or less morally

¹⁰That an act would maximize value may provide a reason for performing it. We cannot assume, however, that the implied reason is or entails a moral prescription.

¹¹It does not matter how we define the class of utilitarian moral principles. The point is, members of that class are distinct from the guiding utilitarian idea.

¹²I shall have some more to say about these assumptions later.

fortunate, about which he is relentlessly direct,¹³ he is free to become indirect. Indirect and intricate any plausible account of, say, moral rightness can be expected to be quite elaborate, involving not only questions about rules or principles, but also about motivations, dispositions to feel guilt or attribute blame, and so on. (Railton 1988, 41f)

The challenge for valoric utilitarianism is to integrate these diverse elements, each of which is based on generic utilitarian calculations, in order to arrive at a plausible utilitarian account of right conduct. A valoric utilitarian account of rightness would be influenced by utilitarian assessments of (for example) the attitudes accompanying actions as well as by utilitarian assessments of the behavior itself.

A familiar example will suggest why, and will also help to relate Railton's project to our problem. Suppose that shortages of food, fuel, or water threaten the general welfare, but that we have access to adequate shares of such necessities because of general compliance with rationing restrictions. Suppose that I am in a position to take more than my share without endangering the rationing arrangement or risking detection. My doing so may be condemned as "free-riding."

Familiar utilitarian moral theories appear inadequate to account for the relevant moral judgments. Actual consequence act utilitarianism implies that, in the circumstances imagined, I should take advantage of others' compliance with the rationing arrangement by free-riding. Probabilistic act utilitarianism might argue either way, depending on the circumstances. If it recommends free-riding, it faces the moral objection that was just levelled against actual-consequence act utilitarianism. If it counsels compliance, it does so for reasons that may be considered morally inadequate, such as the risk of undermining the rationing regime. That factor accounts for part, but only part, of the relevant moral concerns.

That it accounts for only part of the moral concerns can be explained by reference to the possibility that taking more than my share is not wrong. When taking more would not endanger the rationing regime, for example, the morality of my doing so depends on my reasons -- whether I take more because I think I can do so with impunity or because I know that the extra goods are urgently needed by someone (either myself or another person). It may be morally permissible for me to take more than my share in order to meet someone's urgent need. In fact, taking more for that reason in those circumstances would not even count as "free-riding."¹⁴

The rationing example is interesting because we can imagine a generic utilitarian account of at least some of the various moral considerations which must somehow be combined in order to construct a valoric utilitarian account of right conduct. An overall valoric utilitarian appraisal of my action would be based on the utility of my attitudes as well as the utility of my taking more than my share. Some of the relevant utilitarian considerations are familiar. On the one hand, it is

¹³This suggests that the initial stage of appraisal ranks particular items in each class much the way act utilitarianism ranks particular acts. When acts alone are ranked, however, all other conditions are given. It is unclear how to rank all of these particular items when the other conditions that affect a particular item's utility include particular items in other classes whose members are likewise to be ranked and for that reason would not seem to be regarded as given.

¹⁴Note that my comments on this example are not purely utilitarian. For the aim of valoric utilitarianism is not to maximize utility but to use generic utilitarian elements in the construction of a defensible criterion of right conduct. I'll get back to this point in a moment.

generally useful for us to comply with necessary rationing restrictions, but it can be even more useful to take extra benefits in special circumstances in order to meet urgent needs that are not addressed by the restrictions. On the other hand, it is useful for people to have a moral disposition to condemn free-riding and value fair play. That is because the temptation to ride free can be both strong and harmful, and a firm moral aversion is needed to prevent counter-productive free-riding, which by force of example can undermine useful cooperation. However, the most useful aversion would presumably be flexible enough to take due account of urgent needs which are not reflected in the rationing restrictions.

The challenge for valoric utilitarianism is to integrate such diverse considerations. A moral judgment of my action would be determined by a complex principle that appropriately combines a generic utilitarian appraisal of behavior with, for example, a generic utilitarian appraisal of the relevant states of mind.

This incomplete sketch of valoric utilitarianism begins to suggest what the utilitarian moral principle I am seeking might be like. It is important to see, however, that valoric utilitarianism is not that principle.

Railton does **not** ask what kind of utilitarian moral principle most faithfully represents the guiding utilitarian idea. He imagines a utilitarian moral theory that might “overcome some of the difficulties facing” the more familiar utilitarian moral theories. (Railton 1988, p. 410) His concern is the possibility of constructing a morally adequate version of utilitarianism. That project is different from mine, as I shall now explain.

Consider the idea that candidate moral theories should be subjected to a process like the one that is supposed to result in what John Rawls has termed “reflective equilibrium.” (Rawls 1971, p. 48) The process is believed to give reasons for preferring the moral conceptions that emerge from it to those that enter it. We begin with a select set of moral principles and moral judgments at various levels as well as other relevant information. We try to reconcile them all, as far as possible, perhaps modifying some principles or judgments, perhaps rejecting some entirely. The aim is to come up with a set of more defensible principles and judgments.

Imagine that among the moral principles fed into this process is one that best, most faithfully represents the guiding utilitarian idea. The principle may be modified in the process so as to make it a more defensible moral theory. Now we can compare Railton's concerns with ours.

Railton is concerned with a principle that might enjoy reflective equilibrium with other considered moral judgments rather than a principle that would be fed into the equilibrating process. He assumes that utilitarian moral theories can take a variety of forms. Partly to accommodate moral objections to familiar utilitarian principles, he imagines the construction of a complex theory that would integrate various utilitarian factors.

Whereas Railton is concerned with the **output** of the equilibrating process, my concern in this paper is with the **input**. My question is, in effect, **What moral principle should initially represent utilitarianism -- prior, that is, to possibly modifying it so as to bring the theory into reflective equilibrium with other considered moral judgments?**

VII. Multiple-Influence Principles

An answer to that question is suggested by Amartya Sen. In discussing the basic

elements of utilitarianism, Sen identifies a principle which he regards as common to “all variants” of the theory. (Sen 1979, p. 463) He names the principle “outcome utilitarianism” and formulates it as follows:

Any state of affairs x is at least as good as an alternative state of affairs y if and only if the sum total of individual utilities in x is at least as large as the sum total of individual utilities in y . (p. 464)

Sen claims that “A utilitarian moral structure consists of the central element of outcome utilitarianism combined with some consequentialist method of translating judgments of outcomes into judgments of actions.” (p. 466)¹⁵ He depicts act utilitarianism as the result of combining outcome utilitarianism with act consequentialism, which he understands as follows:

An action α is right if and only if the state of affairs x resulting from α is at least as good as each of the alternative states of affairs that would have resulted respectively from each of the alternative acts. (p. 464)

Sen calls acts and rules “influencing variables.” He calls theories such as act utilitarianism and rule utilitarianism “single-influence” utilitarianisms. He distinguishes single-influence from multiple-influence theories, and makes the following claim. He asserts that any single-influence utilitarianism “fails to achieve the best outcome that could have resulted from a comprehensive structure.” (p. 466)¹⁶

I understand this as the generic utilitarian claim that conformity to some (as yet unspecified) utilitarian moral principle, which refers directly to more than one factor -- which refers to, say, rules and behavioral dispositions as well as acts -- would promote welfare better than would conformity to any utilitarian moral principle that refers only to a single factor. For Sen also asserts that

The most comprehensive consequentialist structure would require that the combination of all influencing variables be so chosen that the result is the best feasible state of affairs according to outcome utilitarianism. (p. 466)

Although Sen's argument does not require him to elaborate on that claim, he points in the general direction of what I am seeking.

Railton imagines a moral principle that **least controversially** represents utilitarianism in the moral sphere -- a principle calculated to meet objections against utilitarian moral theories. Sen imagines a moral principle that **most faithfully** represents utilitarianism in the moral sphere -- a principle designed to ensure that conformity to it maximizes welfare.

VIII. The Moral Opacity of Utilitarianism

What might such a principle be like? Although I am not confident about this, I shall assume that there is a principled basis for determining one way of specifying each of the various factors that must be incorporated in a comprehensive utilitarian moral principle for appraising

¹⁵Sen's “utilitarian moral structures” support moral appraisals of acts only, although outcome utilitarianism would not seem to imply any such restriction and utilitarian moral appraisals would seem possible of things other than acts.

¹⁶Although Sen refers to an argument establishing this point (Sen 1979, 466n8), I have not found it.

conduct (each of what Sen calls “influencing variables”).

Consider the case of rules. We might select them on the basis of their actual or their predictable consequences. Consequences of either type might be attributed to rules on the basis of perfect conformity or some more realistic criterion. For the guiding utilitarian idea to have a best moral representative, the principle's reference to rules (as well as to other variables) must be unambiguous. Perhaps the selection criteria for the rules should be determined so as to maximize the utility of the resulting principle.

Another thing that is unclear is how such a principle integrates the various factors so as to generate authoritative utilitarian moral appraisals of particular actions. Call this the combination problem.¹⁷ This problem indicates that utilitarianism is morally opaque. There may be a unique principle of right conduct that best represents utilitarianism in the moral sphere, but we have no idea what it might be like.¹⁸

The combination problem may be explained informally as follows. It is unclear how much of our efforts should be dedicated to producing the best consequences by our actions and how much should be diverted (when a choice seems needed) to promoting useful institutions, laws, social rules and practices, character and personality traits, and so on, and how much one's personal division of labor should vary with circumstances and how such a division of labor should vary from one individual to another.

The free-riding example might appear to suggest at least a partial solution to the combination problem. As I have interpreted ordinary moral judgment, it implies that the morality of taking more than one's fair share is determined largely by the agent's attitudes. This suggests a generalization, namely, that the natural expressions in conduct of useful dispositions should take precedence over alternative courses of action with greater marginal utility.

But we cannot generalize in that way if we wish to construct a utilitarian moral principle that best represents the guiding utilitarian idea. The suggested precedence derives from a considered moral judgment rather than generic utilitarian reasoning. To solve the combination problem, we need a principle conformity to which would maximize welfare, and we cannot assume that such a principle would agree with our considered moral judgments and thus would generally condemn free-riding or favor the natural expressions in conduct of useful dispositions over alternative courses of action with greater marginal utility.

To get a rough idea of the sort of principle we seek as well as the difficulties involved, we might begin with an artificially simple principle that takes into account the utility of rules as well as acts. Generic utilitarianism would seem to argue for a principle that calls for compliance with useful rules, except when welfare would be promoted better in the long run by deviation from them, and to call for conduct that maximizes welfare, except when it would be more useful in the long run for conduct to comply with the useful rules.

Application of such a principle must determine when it is useful to promote useful rules, what useful rules should be promoted in what circumstances, what effort to expend on that

¹⁷We could avoid this problem by applying the influencing variables and balancing them in an “intuitionistic” manner. This would not yield a determinate principle, but it suggests a model for utilitarian reasoning in practice.

¹⁸I think it is also unclear how to determine whether a principle best or most faithfully expresses the guiding utilitarian idea. Call that the representation problem. I'll come back to it in the next section.

project, what effort to expend on promoting compliance with actual rules, and what effort to expend on promoting compliance with possible rules. As this incomplete catalogue of questions implies, both the identification and the application of such a principle are complicated by the fact that various rules, with varying degrees of utility, actually exist within the various social groups whose rules apply to one's conduct, and that such rules are actually complied with to a greater or lesser degree.

There might very well be a determinate answer to such questions, at least in theory. But as the relevant facts are largely unknowable, we seem bound to remain ignorant of how to frame such a principle. We have no reason to expect it ever to become available for use as a criterion of right conduct or as the basis of a decision procedure.

As we have seen, there are generic utilitarian reasons for the principle we are seeking to include direct consideration of other factors, such as useful dispositions. Incorporating them can be expected to complicate matters further.

I'll now suggest what a principle with this third factor added would involve. Generic utilitarianism would seem to argue for a principle that calls for conduct that maximizes welfare, except when it would be more useful in the long run for conduct to express useful dispositions or to comply with useful rules. The principle would not permit useful conduct to undermine dispositions or rules that it would be more useful in the long run to maintain, though the latter would presumably vary with circumstances. It would require that conduct promote useful dispositions and useful rules, but only when it would be more useful in the long run to do so.

I assume that we can theoretically state the implications for action of behavioral dispositions as readily as we can for useful rules. Rules proclaim their requirements for conduct. Behavioral dispositions have natural expressions in conduct. We sometimes act at variance with our stable dispositions, and utility might sometimes be maximized by so acting. It will be the task of the principle we seek to take account of that complication.

Application of such a principle must determine not only when it is useful to promote useful dispositions, but also what dispositions they might be, what effort to expend on that project, and when we should promote useful rules, useful dispositions, or more directly useful actions. This will be made somewhat more complex by the fact that individuals at any time have various dispositions, with varying degrees of utility, and varying degrees of firmness.

Some of the problems that accrue to identifying useful rules, for the purpose of applying the principle, accrue as well to identifying useful behavioral dispositions. The utility of behavioral dispositions, like that of rules, varies with circumstances. A complicating factor is that we must consider among the circumstances other characteristics of the individuals whose dispositions they are, for dispositions are likely to interact.

I do not believe we are at all likely to have reliable information that would make possible a more practically helpful formulation of such a principle. This means that such a principle will not be available for use as a criterion of right conduct or as the basis of a decision procedure.

The combination problem suggests that, if there is a most faithful moral representative of utilitarianism, it has implications for conduct that are practically unknowable. I have assumed that the variables that must be incorporated into a comprehensive utilitarian moral principle have determinate values that are fixed somehow by the guiding utilitarian idea. I am not confident that this is the case. If it is not the case, then my argument suggests that there is no best moral

interpretation of the guiding utilitarian idea.

IX. Moral Ambiguity

I come finally to the representation problem. This concerns the criteria for determining what moral principle most faithfully represents utilitarianism. My argument so far has taken for granted that this is no problem at all. It has assumed that a moral principle most faithfully represents the guiding utilitarian idea if conformity to it would most effectively promote welfare. I shall now question that assumption.

We can raise at least three questions about the test which I have supposed is implied by the guiding utilitarian idea: (1) One principle for everyone, or more than one? (2) One principle for each person, or more than one? and (3) What constitutes the relevant kind of conformity?

(1) The first question is whether, in comparing the effects of conformity to different principles, we are simply to compare the effects of conformity to each principle, in turn, by everyone. For there is an alternative, which is to compare the effects of conformity to different principles by different individuals. It would seem advisable to contemplate this possibility because it is conceivable that welfare would be better promoted if different individuals conformed to different principles.¹⁹ Utilitarianism would seem to provide no way of deciding between these two different ways of comparing utilitarian moral principles.²⁰

(2) The second question is whether, in comparing the effects of conformity to different principles, we are simply to compare the effects of continuing conformity to each principle in turn. There is an alternative, however, which is to compare the effects of conformity to a succession of different principles. It would seem advisable to contemplate this possibility because it is conceivable that persons or circumstances might change over time in such a way that welfare would be better promoted if the principles people conformed to varied over time.²¹ Once again, utilitarianism would seem to provide no way of deciding between these two different ways of comparing utilitarian moral principles.²²

(3) My argument has so far assumed that conformity to a principle is to be understood as doing only what the principle permits. Instead of considering the effects of simple conformity to

¹⁹Here's a further complication. The range of principles that we imagine different individuals conforming to might be limited to ones that are recognizably utilitarian; alternatively, we might include within the range of principles some that are not recognizably utilitarian.

²⁰It might be objected that morality requires its principles to apply universally. The alternative criterion does not deny this. It accords with the possibility of fundamental moral disagreement, which the concept of morality does not deny.

²¹The previous complication might apply here. The range of principles that we imagine individuals conforming to might either be limited to ones that are recognizably utilitarian or might include some that are not recognizably utilitarian.

²²It might be objected that morality requires its principles to apply constantly. The alternative criterion does not deny this. It accords with the possibility that individuals change their basic moral commitments over time, which the concept of morality does not deny.

a principle, however, we might consider the effects of its acceptance by individuals. Acceptance involves consistent, conscientious effort to follow a principle but also allows for some deviation from the principle on occasion.²³ It is conceivable that the possible disutility of divergence from a presumably useful principle can sometimes be outweighed by some special utility that might attach to the relaxation of attitudes that acceptance without perfect conformity involves.²⁴

The representation problem, then, is this. If generic utilitarianism determines one single set of criteria for identifying its most faithful representative in the moral sphere, we might reasonably assume that one and only one moral principle best satisfies those criteria. If, however, there are alternative possible ways of identifying utilitarianism's moral representative, we have that much less reason to assume there is just one such principle. In that case, utilitarianism may well be morally ambiguous: substantively different principles of right conduct might equally well qualify for the position of utilitarianism's moral representative, and different principles can be presumed to have differing implications for right conduct.²⁵

If, for any reason, such as the reasons I have suggested, the guiding utilitarian idea fails to determine a unique criterion of right action, it is possible that it might do so with the aid of substantive moral judgments that we are prepared to make without the conscious deployment of utilitarian ideas. To shape a utilitarian moral principle on that basis, however, would be to discard the project of discovering the most faithful moral interpretation of utilitarianism itself and to seek instead a **morally acceptable** version of utilitarianism. In that case, and to that extent, utilitarian moral theory would be driven by external factors rather than by its own guiding utilitarian idea.

REFERENCES

- Adams, Robert M., 1976, "Motive Utilitarianism," *JOURNAL OF PHILOSOPHY*, vol. 73, pp. 1-19.
- Gibbard, Allan F., 1965, "Rule-Utilitarianism: Merely an Illusory Alternative?," *AUSTRALASIAN JOURNAL OF PHILOSOPHY*, vol. 43, pp. 211-220.
- Railton, Peter, 1984, "Alienation, Consequentialism, and the Demands of Morality," *PHILOSOPHY & PUBLIC AFFAIRS*, vol. 13, pp. 134-171.

²³Because acceptance involves certain moral attitudes, whereas simple conformity does not, acceptance may also involve further distinguishing consequences.

²⁴This alternative test would also seem to reflect a more realistic notion of the role we might ideally expect moral principles to play in our lives than perfect conformity does.

²⁵It is also conceivable that some neutral factor, such as the concept of morality or of rationality, would, when added to generic utilitarianism, somehow disqualify all but one potential moral representative of the guiding utilitarian idea and would thus save utilitarianism from moral ambiguity. But I see no reason to assume this.

_____, 1988, "How Thinking about Character and Utilitarianism Might Lead to Rethinking the Character of Utilitarianism," *MIDWEST STUDIES IN PHILOSOPHY*, vol. 13, pp. 398-416.

Rawls, John, 1971, *A THEORY OF JUSTICE*, Cambridge, Mass.: Harvard University Press.

Regan, Donald H., 1980, *UTILITARIANISM AND CO-OPERATION*, New York: Oxford University Press.

Sen, Amartya, 1979, "Utilitarianism and Welfarism," *JOURNAL OF PHILOSOPHY*, vol. 76, pp. 463-489.