MOTIVE UTILITARIANISM *

PHILOSOPHERS have written much about the morality of traits of character, much more about the morality of actions, and much less about the morality of motives. [By "motives" here I mean principally wants and desires, considered as giving rise, or tending to give rise to actions. A desire, if strong, stable, and for a fairly general object (e.g., the desire to get as much money as possible), may perhaps constitute a trait of character; but motives are not in general the same, and may not be as persistent, as traits of character.] Utilitarian theories form a good place to begin an investigation of the relation between the ethics of motives and the ethics of actions, because they have a clear structure and provide us with familiar and comprehensible, if not always plausible, grounds of argument. I believe that a study of possible treatments of motives in utilitarianism will also shed light on some of the difficulties surrounding the attempt to make the maximization of utility the guiding interest of ethical theory.

I

What would be the motives of a person morally perfect by utilitarian standards? It is natural to suppose that he or she would be completely controlled, if not exclusively moved, by the desire to maximize utility. Isn't this ideal of singlemindedly optimific motivation demanded by the principle of utility, if the principle, as Bentham puts it, "states the greatest happiness of all those whose interest is in question, as being the right and proper, and only right and proper and universally desirable, end of human action"? [1]

But there is a good utilitarian objection to such singlemindedness: it is not in general conducive to human happiness. As Sidgwick says, "Happiness [general as well as individual] is likely to be better attained if the extent to which we set ourselves consciously to aim at it be carefully restricted." [2] Suggestions of a utilitarian theory about motivation that accommodates this objection can be found in both Bentham and Sidgwick.

The test of utility is used in different theories to evaluate differ-

* The largest part of my work on this paper was supported by a fellowship from the National Endowment for the Humanities. I am indebted to several, and especially to Gregory Kavka, Jan Narveson, and Derek Parfit, for helpful discussion and comments on earlier versions.

[1] Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (New York: Hafner, 1961) (referred to hereafter as *Introduction,* with page number), p. 1*n*.

[2] Henry Sidgwick, *The Methods of Ethics,* seventh edition (New York: Dover, 1966) (referred to hereafter as *Methods,* with page number), p. 405.

ent objects. It is applied to acts in act utilitarianism and to roles, practices, and types of action in the various forms of rule utilitarianism. In the view about motives stated in the first paragraph above, the test is not applied at all: nothing is evaluated for its utility, but perfect motivation is identified with an all-controlling desire to maximize utility. The test of utility could be applied in various ways in the evaluation of motives.

It could be applied directly to the motives themselves, and is so applied by Bentham, when he says,

> If they [motives] are good or bad, it is only on account of their effects: good, on account of their tendency to produce pleasure, or avert pain: bad, on account of their tendency to produce pain, or avert pleasure (*Introduction,* 102).

Alternatively, we could apply the test directly to objects of desire and only indirectly to the desires, saying that the best motives are desires for the objects that have most utility. Sidgwick seems to take this line when he says,

> While yet if we ask for a final criterion of the comparative value of the different objects of men's enthusiastic pursuit, and of the limits within which each may legitimately engross the attention of mankind, we shall none the less conceive it to depend upon the degree in which they respectively conduce to Happiness (*Methods,* 406).

Or we could apply the test of utility to the acts to which motives give rise (or are likely to give rise) and, thence, indirectly to the motives; the best motives would be those productive of utility-maximizing acts.[3]

Another approach, also endorsed by Bentham, is to evaluate motives by the intentions to which they give rise: "A motive is good, when the intention it gives birth to is a good one; bad, when the intention is a bad one" (*Introduction,* 120). The value of an intention to do an act, he regards as depending, in turn, on whether "the consequences of the act, had they proved what to the agent they seemed likely to be, *would* have been of a beneficial nature" or the opposite (*Introduction,* 93). This approach seems inconsistent with Bentham's insistence that the test of utility must be applied to everything that is to be evaluated—that

> Strictly speaking, nothing can be said to be good or bad, but either in itself; which is the case only with pain or pleasure: or on account

[3] This too may find some support in Sidgwick. Cf. *Methods,* 493, on the praise of motives conceived to prompt to felicific conduct.

of its effects; which is the case only with things that are the causes or preventives of pain and pleasure (*Introduction,* 87; cf. 102).

Bentham would presumably defend the evaluating of intentions by the utility of expected consequences of the intended act rather than the utility of the intentions themselves in the same way that he defends a similar method of evaluating dispositions. That is, he would appeal to the assumption "that in the ordinary course of things the consequences of actions commonly turn out conformable to intentions" (*Introduction,* 133), so that there is no practical difference between the utility of the intention and the utility of the expected consequences of the intended action. This assumption is plausible as regards the short-term consequences of our actions, though even there it yields at best a very rough equivalence between utility of intentions and utility of expected consequences. It is wildly and implausibly optimistic as regards our ability to foresee the long-term consequences of our actions.[4]

Bentham similarly regards the evaluating of motives by the value of intentions arising from them as consistent with (or even practically equivalent to) a direct application of the test of utility to motives, on the ground that the intention resulting from a motive is responsible for "the most material part of [the motive's] effects" (*Introduction,* 120). His position will still be inconsistent, however, unless he maintains (falsely, I believe) that the resulting intentions to act are responsible for *all* the relevant effects of having a motive.

If the moral point of view, the point of view from which moral evaluations are made, is dominated by concern for the maximization of human happiness, then it seems we must revert to the thesis that the test of utility is to be applied directly to everything, including motives. This is the conclusion toward which the following argument from Sidgwick tends:

> Finally, the doctrine that Universal Happiness is the ultimate *standard* must not be understood to imply that Universal Benevolence is the only right or always best *motive* of action. For . . . if experience shows that the general happiness will be more satisfactorily attained if men frequently act from other motives than pure universal philanthropy, it is obvious that these other motives are reasonably to be preferred on Utilitarian principles (*Methods,* 413).

[4] Also, as Gregory Kavka has pointed out to me, the utility of *having* an intention (e.g., to retaliate if attacked) may be quite different from the utility (actual or expected) of *acting* on it. I shall be making a similar point about motives, below.

Accordingly, the theory that will be my principal subject here is that one pattern of motivation is morally better than another to the extent that the former has more utility than the latter. The morally perfect person, on this view, would have the most useful desires, and have them in exactly the most useful strengths; he or she would have the most useful among the patterns of motivation that are causally possible for human beings.[5] Let us call this doctrine *motive utilitarianism.*

II

It is distinct, both theoretically and practically, from act utilitarianism. It can be better, by motive-utilitarian standards, to have a pattern of motivation that will lead one to act wrongly, by act-utilitarian standards, than to have a motivation that would lead to right action. Even if there is no difference in external circumstances, the motivational pattern that leads to more useful actions is not necessarily the more useful of two motivational patterns, on the whole. For the consequences of any acts one is thereby led to perform are not always the only utility-bearing consequences of being influenced, to a given degree, by a motive.[6]

This can be seen in the following fictitious case. Jack is a lover of art who is visiting the cathedral at Chartres for the first time. He is greatly excited by it, enjoying it enormously, and acquiring memories which will give him pleasure for years to come. He is so excited that he is spending much more time at Chartres than he had planned, looking at the cathedral from as many interior and exterior angles, and examining as many of its details, as he can. In fact, he is spending too much time there, from a utilitarian point of view. He had planned to spend only the morning, but he is spending the whole day; and this is going to cause him considerable inconvenience and unpleasantness. He will miss his dinner, do several hours of night driving, which he hates, and have trouble finding a place to sleep. On the whole, he will count the day well spent, but some of the time spent in the cathedral will not produce as much utility as would have been produced by departing that

[5] It is difficult to say what is meant by the question, whether a certain pattern of motivation is causally possible for human beings, and how one would answer it. I shall sidestep these issues here, for I shall be making comparative evaluations of motives assumed to be possible, rather than trying to determine the most useful of all causally possible motivations.

[6] I am here denying, as applied to motives, what Bernard Williams rather obscurely calls the "act-adequacy premise" [A Critique of Utilitarianism," in J. J. C. Smart and Williams, *Utilitarianism, For and Against* (New York: Cambridge, 1975), pp. 119–130].

much earlier. At the moment, for example, Jack is studying the sixteenth to eighteenth century sculpture on the stone choir screen. He is enjoying this less than other parts of the cathedral, and will not remember it very well. It is not completely unrewarding, but he would have more happiness on balance if he passed by these carvings and saved the time for an earlier departure. Jack knows all this, although it is knowledge to which he is not paying much attention. He brushes it aside and goes on looking at the choir screen because he is more strongly interested in seeing, as nearly as possible, everything in the cathedral than in maximizing utility. This action of his is therefore wrong by act-utilitarian standards, and in some measure intentionally so. And this is not the only such case. In the course of the day he knowingly does, for the same reason, several other things that have the same sort of act-utilitarian wrongness.

On the other hand, Jack would not have omitted these things unless he had been less interested in seeing everything in the cathedral than in maximizing utility. And it is plausible to suppose that if his motivation had been different in that respect, he would have enjoyed the cathedral much less. It may very well be that his caring more about seeing the cathedral than about maximizing utility has augmented utility, through enhancing his enjoyment, by more than it has diminished utility through leading him to spend too much time at Chartres. In this case his motivation is right by motive-utilitarian standards, even though it causes him to do several things that are wrong by act-utilitarian standards.

Perhaps it will be objected that the motive utilitarian should say that Jack ought indeed to have been as interested in the cathedral as he was, but ought to have been even more interested in maximizing utility. Thus he would have had as much enjoyment from the more rewarding parts of the cathedral, according to the objector, but would not have spent too much time on the less rewarding parts. The weak point in this objection is the assumption that Jack's enjoyment of the things he would still have seen would not be diminished in these circumstances. I think, and I take it that Sidgwick thought too,[7] that a great concern to squeeze out the last drop of utility is likely to be a great impediment to the enjoyment of life. Therefore it seems plausible to suppose that from a motive-utilitarian point of view Jack ought not only to have been as

[7] I believe this is the most natural reading of Sidgwick, but it may be barely possible to construe him as meaning only that the perpetual *consciousness* of such a concern would be an impediment. See *Methods*, 48f.

strongly interested in seeing the cathedral as he was, but also to have been as weakly interested in maximizing utility as he was.

In describing this case I have been treating the maximization of utility as a unitary end which Jack might have pursued for its own sake. Perhaps it will be suggested that, although an all-controlling desire for that end would have diminished utility by dulling Jack's enjoyment, he could have had undimmed enjoyment without wrong action if he had had the maximization of utility as an *inclusive end*—that is, if he had been moved by desire for more particular ends for their own sakes, but in exact proportion to their utility.[8] But this suggestion is not plausible. While he is in the cathedral Jack's desire to see everything in it is stronger, and his desire for the benefits of an early departure is weaker, than would be proportionate to the utility of those ends. And a stronger desire for an early departure would probably have interfered with his enjoyment just as much as a stronger desire for utility maximization as such. We are likely in general to enjoy life more if we are often more interested in the object of an enthusiastic pursuit, and less concerned about other ends, than would be proportionate to their utility. It follows that failing (to some extent) to have utility maximization as an inclusive end is often right by motive-utilitarian standards, and may be supposed to be so in Jack's case.

In order to justify the view that motive utilitarianism implies something practically equivalent to act utilitarianism one would have to show that the benefits that justify Jack's motivation by motive-utilitarian standards also justify his spending time on the choir screen by act-utilitarian standards. But they do not. For they are not consequences of his spending time there, but independent consequences of something that caused, or manifested itself in, his spending time there. It is not that deciding to devote only a cursory inspection to the choir screen would have put him in the wrong frame of mind for enjoying the visit. It is rather that, being in the right frame of mind for enjoying the visit, he could not bring himself to leave the choir screen as quickly as would have maximized utility.

III

The act utilitarian may try to domesticate motive utilitarianism, arguing (A) that motive utilitarianism is merely a theorem of act

[8] The terminology of "dominant" and "inclusive" ends was developed by W. F. R. Hardie, "The Final Good in Aristotle's Ethics," *Philosophy*, XL, 154 (October 1965): 277–295; Rawls makes use of it. J. S. Mill seems to treat the maximization of utility as an inclusive end in *Utilitarianism*, ch. 4, §§ 5–8.

utilitarianism, and denying (B) that behavior like Jack's inspection of the choir screen, if resulting from obedience to the dictates of motive utilitarianism, can properly be called wrong action.

(A) Since act utilitarianism implies that one ought to do whatever has most utility, it implies that, other things equal, one ought to foster and promote in oneself those motives which have most utility. And that, it may be claimed, is precisely what motive utilitarianism teaches.

(B) Jack was once, let us suppose, an excessively conscientious act utilitarian. Recognizing the duty of cultivating more useful motives in himself, he took a course of capriciousness training, with the result that he now stands, careless of utility, before the choir screen. It would be unfair, it may be argued, to regard what Jack is now doing as a wrong action by utilitarian standards. Rather, we must see it as only an inescapable part of a larger, right action, which began with his enrolling for capriciousness training—just as we do not say that a person rightly jumped from a burning building, saving his life, but wrongly struck the ground, breaking his leg. It is unreasonable, on this view, to separate, for moral evaluation, actions that are causally inseparable.

Both of these arguments are to be rejected. The second (B) involves deep issues about the individuation of actions and the relation between causal determination and moral responsibility. It seems clear enough, however, that Jack's staying at the choir screen is separable from his earlier efforts at character reform in a way that striking the ground is not separable from jumping out of a building. Once you have jumped, it is no longer in your power to refrain from striking the ground, even if you want to. If you are sane and well informed about the situation, you have only one choice to make: to jump or not to jump. There is no further choice about hitting the ground, and therefore it is inappropriate to separate the impact from the leap, as an object of moral evaluation. But even after Jack has taken capriciousness training, it is still in his power to leave the choir screen if he wants to; it is just that he does not want to. His choice to stay and examine it is a new choice, which he did not make, years ago, when he decided to reform. He did decide then to become such that he would sometimes make nonutilitarian choices, but it may not even have occurred to him then that he would ever be in Chartres. It seems perfectly appropriate to ask whether the choice that he now makes is morally right or wrong.

It is plausible, indeed, to say that Jack is not acting wrongly in acting on the motivation that he has rightly cultivated in himself. But I think that is because it is plausible to depart from act utilitarianism at least so far as to allow the rightness or wrongness of Jack's action in this case to depend partly on the goodness or badness of his motive, and not solely on the utility of the act. It is noteworthy in this connection that it would be no less plausible to acquit Jack of wrongdoing if he had always been as easygoing as he now is about small increments of utility, even though there would not in that case be any larger action of character reform, of which Jack's present scrutiny of the choir screen could be regarded as an inescapable part.

A similar irrelevant emphasis on doing something about one's own motivational patterns also infects the attempt (A) to derive motive utilitarianism from act utilitarianism. Motive utilitarianism is not a theorem of act utilitarianism, for the simple reason that motive utilitarianism is not about what motives one ought to foster and promote, or *try* to have, but about what motives one ought to *have*. There is a preconception to be overcome here which threatens to frustrate from the outset the development of any independent ethics of motives. I refer to the assumption that "What should I (try to) do?" is *the* ethical question, and that we are engaged in substantive *ethical* thinking only insofar as we are considering *action*-guiding principles.[9] If we hold this assumption, we are almost bound to read "What motives should I have?" as "What motives should I try to develop and maintain in myself?"

There are other questions, however, that are as fundamental to ethics as "What should I do?" It is characteristic of moral as opposed to pragmatic thinking that, for example, the question, "Have I lived well?" is of interest for its own sake. In pragmatic self-appraisal that question is of interest only insofar as the answer may guide me toward future successes. If I am personally concerned, in more than this instrumental way, and not just in curiosity, about whether I have lived well, my concern is not purely pragmatic, but involves at least a sense of style, if not of morality.

If the question is "Have I lived well?" the motives I have *had* are relevant, and not just the motives I have *tried* to have. If I

[9] Cf. Jan Narveson, *Morality and Utility* (Baltimore, Md.: John Hopkins, 1967), p. 105: "Let us begin by recalling the primary function of ethical principles: to tell us what to do, i.e., to guide action. Whatever else an ethical principle is supposed to do, it must do that, otherwise it could not (logically) be an ethical principle at all."

tried to have the right motive, but nonetheless had the wrong one —if I tried to love righteousness and my neighbors, but failed and did my duty out of fear of hellfire for the most part—then I did not live as well as I would have lived if I had *had* the right motive.

Suppose, similarly, that Martha is an overscrupulous utilitarian, completely dominated by the desire to maximize utility. She has acted rightly, by act-utilitarian standards, just as often as she could. Among her right actions (or attempts at right action) are many *attempts* to become strongly interested in particular objects—more strongly, indeed, than is proportionate to their utility. For she realizes that she and her acquaintances would be happier if she had such interests. But all these attempts have failed.

Mary, on the other hand, has not had to work on herself to develop such nonutilitarian interests, but has always had them; and, largely because of them, her motivational patterns have had more utility, on the whole, than Martha's. The motive utilitarian will take this as a reason (not necessarily decisive) for saying that Martha has *lived less well* than Mary. This censure of Martha's motives is not derivable from act utilitarianism, for her actions have been the best that were causally possible for her. (If you are tempted to say that Martha's conscientiousness is better than Mary's more useful motives, you are experiencing a reluctance to apply the test of utility to motives.)

IV

I have argued that right action, by act-utilitarian standards, and right motivation, by motive-utilitarian standards, are incompatable in some cases. It does not immediately follow, but it may further be argued, that act utilitarianism and motive utilitarianism are incompatible theories.

One argument for this conclusion is suggested, in effect, by Bernard Williams. He does not formulate or discuss motive utilitarianism, but he holds that it is inconsistent of J. J. C. Smart, following Sidgwick, "to present direct [i.e., act] utilitarianism as a doctrine merely about justification and not about motivation." Williams's argument is,

> There is no distinctive place for *direct* utilitarianism unless it is, within fairly narrow limits, a doctrine about how one should decide what to do. This is because its distinctive doctrine is about what acts are right, and, especially for utilitarians, the only distinctive interest or point of the question what acts are right, relates to the situation of deciding to do them (*op. cit.*, 128).

The doctrine about motives that Williams believes to be implied by act utilitarianism is presumably the doctrine, discarded at the beginning of my present essay, that one ought always to be controlled by the desire or purpose of maximizing utility. And this doctrine, if conjoined with plausible empirical beliefs illustrated in section II above, is inconsistent with motive utilitarianism.

There are two questionable points in Williams's argument. One is the claim that for utilitarians the only use of the question, What acts are right? is for guidance in deciding what to do. He defends this claim, arguing that "utilitarians in fact are not very keen on people blaming themselves, which they see as an unproductive activity," and that they therefore will not be interested in the question, "Did he (or I) do the right thing?" (124). I am not convinced by this defense. Blame is a self-administered negative reinforcement which may perhaps cause desirable modifications of future behavior. The retrospective question about the evaluation of one's action is a question in which one can hardly help taking an interest if one has a conscience; one who desires to act well will naturally desire to *have* acted well. And the desire to act well, at least in weighty matters, will surely be approved on motive-utilitarian grounds.

But suppose, for the sake of argument, we grant Williams that the point of act-utilitarian judgments, when they have a point, is to guide us in deciding what to do. His argument still rests on the assumption that the act utilitarian is committed to the view that it is generally useful to ask what acts are right, and that one ought always or almost always to be interested in the question. Why should the act utilitarian be committed to this view? If he is also a motive utilitarian, he will have reason to say that, although it is indeed useful to be guided by utilitarian judgments in actions of great consequence, it is sometimes better to be relatively uninterested in considerations of utility (and so of morality). "For everything there is a season and a time for every matter under heaven: . . . a time to kill, and a time to heal; a time to break down, and a time to build up," said the Preacher (Ecclesiastes 3:1, 3 RSV). The act-and-motive utilitarian adds, "There is a time to be moral, and a time to be amoral." (The act-and-motive utilitarian is one who holds both act and motive utilitarianism as *theories*. He does not, for he cannot, always satisfy the demands of both theories in his acts and motives.)

Perhaps it will be objected that this reply to Williams overlooks the utility of conscientiousness. Conscience is, in part, a motive:

the desire to act or live in accordance with moral principles. If the moral principles are mainly sound, it is so useful a motive that it is important, from a motive-utilitarian standpoint, not to undermine it. This consideration might make a motive utilitarian reluctant to approve the idea of "a time to be amoral," lest such "moral holidays" weaken a predominantly useful conscience.

The question facing the act-and-motive utilitarian at this point is, what sort of conscience has greatest utility. We have seen reason to believe that an act-utilitarian conscience that is scrupulous about small increments of utility would have bad effects on human happiness, smothering many innocent enjoyments in a wet blanket of excessive earnestness. A more useful sort of conscience is probably available to the act-and-motive utilitarian. It would incorporate a vigorous desire to *live well,* in terms of the over-all utility of his life, but not necessarily to *act rightly* on every occasion. Having such a conscience, he would be strongly concerned (1) not to act in ways gravely detrimental to utility, and (2) not to be in a bad motivational state. If he performs a mildly unutilitarian action as an inevitable consequence of the most useful motivation that he can have, on the other hand, he is still living as well as possible, by his over-all utilitarian standards; and there is no reason why such action should undermine his determination to live well. A conscience of this sort seems as possible, and at least as likely to be stable, as a conscience that insists on maximizing utility in every action. Thus the act-and-motive utilitarian has good motive-utilitarian reasons for believing that he should sometimes be, in relation to his act-utilitarian principles, amoral.

V

But this conclusion may be taken, quite apart from Williams's argument, as grounds for thinking that act utilitarianism and motive utilitarianism are incompatible in the sense that holding the latter ought reasonably to prevent us from holding the former as a *moral* theory. The incompatibility has to do with moral seriousness. The problem is not just that one cannot *succeed* in living up to the ideals of both theories simultaneously. It is rather that the motive utilitarian is led to the conclusion that it is morally better on many occasions to be so motivated that one will not even *try* to do what one ought, by act-utilitarian standards, to do. If the act-and-motive utilitarian accepts this conclusion, however, we must wonder whether all his act-utilitarian judgments about what one ought to do are really judgments of *moral* obligation. For it is commonly made a criterion for a theory's being a theory of *moral*

obligation, that it claim a special seriousness for its judgments of obligation. By this criterion, act utilitarianism cannot really be a theory of moral obligation (as it purports to be) if it is conjoined with the view that some of its dictates should be taken as lightly as motive utilitarianism would lead us to think they should be taken.

This argument depends on the triviality of any reasonable human interest in some of the obligations that act utilitarianism would lay on us. And the triviality is due to the totalitarian character of act utilitarianism, to its insistence that, as Sidgwick puts it, "it is *always* wrong for a man knowingly to do *anything* other than what he believes to be most conducive to Universal Happiness" (*Methods,* 492, italics mine).

Without this triviality a conflict between the ethics of actions and the ethics of motives need not destroy the seriousness of either. Maybe *no* plausible comprehensive ethical theory can avoid all such conflicts. Are there *some* circumstances in which it is best, for example, in the true morality of motives, to be unable to bring oneself to sacrifice the happiness of a friend when an important duty obliges one, in the true morality of actions, to do so? I don't know. But if there are, the interests involved, on both sides, are far from trivial, and the seriousness of both moralities can be maintained. If one fails to perform the important duty, one ought, seriously, to feel guilty; but one could not do one's duty in such a case without having a motivation of which one ought, seriously, to be ashamed. The situation presents a tragic inevitability of moral disgrace.

There are, accordingly, two ways in which the utilitarian might deal with the argument if he has been trying to combine act and motive utilitarianism and accepts the view I have urged on him about the kind of conscience it would be most useful to have. (A) He could simply acknowledge that he is operating with a modified conception of moral obligation, under which a special seriousness attaches to some but not all moral obligations.[10] He would claim that his use of "morally ought" nonetheless has enough similarity, in other respects, to the traditional use, to be a reasonable extension of it.

(B) The other, to my mind more attractive, way is to modify the act-utilitarian principle, eliminating trivial obligations, and limit-

[10] It may be thought that Sidgwick has already begun this modification, by holding that good actions ought not to be praised, nor bad ones blamed, except insofar as it is useful to praise and blame them. See *Methods,* 428 f., 493.

ing the realm of duty to actions that would be of concern to a conscience of the most useful sort. Under such a limitation it would not be regarded as morally wrong, in general, to fail to maximize utility by a *small* margin. One's relatively uninfluential practical choices would be subject to moral judgment only indirectly, through the motive-utilitarian judgment on the motives on which one acted (and perhaps a character-utilitarian judgment on the traits of character manifested by the action). Some acts, however, such as shoplifting in a dime store or telling inconsequential lies, would still be regarded as wrong even if only slightly detrimental in the particular case, because it is clear that they would be opposed by the most useful sort of conscience. I leave unanswered here the question whether a conscience of the most useful kind would be offended by some acts that maximize utility—particularly by some utility-maximizing violations of such rules as those against stealing and lying. If the answer is affirmative, the position we are considering would have approximately the same practical consequences as are commonly expected from rule utilitarianism. This position—that we have a *moral duty* to do an act, if and only if it would be demanded of us by the most useful kind of conscience we could have—may be called "conscience utilitarianism," and is a very natural position for a motive utilitarian to take in the ethics of actions.

The moral point of view—the point of view from which moral judgments are made—cannot safely be defined as a point of view in which the test of utility is applied directly to all objects of moral evaluation. For it is doubtful that the most useful motives, and the most useful sort of conscience, are related to the most useful acts in the way that the motives, and especially the kind of conscience, regarded as right must be related to the acts regarded as right in anything that is to count as a morality. And therefore it is doubtful that direct application of the test of utility to everything results in a system that counts as a morality.

VI

Considered on its own merits, as a theory in the ethics of motives, which may or may not be combined with some other type of utilitarianism in the ethics of actions, how plausible is motive utilitarianism? That is a question which we can hardly begin to explore in a brief paper, because of the variety of forms that the theory might assume, and the difficulty of stating some of them. The exploration might start with a distinction between individualistic

and universalistic motive utilitarianism, analogous to the distinction between act and rule utilitarianism.

Individualistic motive utilitarianism holds that a person's motivation on any given occasion is better, the greater the utility of *his* having it on *that* occasion. This seemed to Bentham, on the whole, the least unsatisfactory view about the moral worth of motives:

> The only way, it should seem, in which a motive can with safety and propriety be styled good or bad, is with reference to its effects *in each individual instance* (*Introduction,* 120, italics mine).

This doctrine seems liable to counterexamples similar to those which are commonly urged against act utilitarianism. An industrialist's greed, a general's bloodthirstiness, may on some occasions have better consequences on the whole than kinder motives would, and even predictably so. But we want to say that they remain worse motives.

Universalistic motive utilitarianism is supposed to let us say this, but is difficult to formulate. If we try to state it as the thesis that motives are better, the greater the utility of *everybody's* having them on *all* occasions, we implausibly ignore the utility of diversity in motives. A more satisfactory view might be that a motivation is better, the greater the average probable utility of *anyone's* having it on *any* occasion. This formulation gives rise to questions about averaging: do we weigh equally the utility of a motive on all the occasions when it could conceivably occur, or do we have some formula for weighing more heavily the occasions when it is more likely to occur? There are also difficult issues about the relevant description of the motive. One and the same concrete individual motive might be described correctly as a desire to protect Henry Franklin, a desire to protect (an individual whom one knows to be) one's spouse, a desire to protect (an individual whom one knows to be) the chief executive of one's government, and a desire to protect (an individual whom one knows to be) a betrayer of the public trust; these motive types surely have very different average utilities. If one makes the relevant description of the motive too full, of course, one risks making universalistic motive utilitarianism equivalent to individualistic.[11] If the description is not full enough, it will be hard to get any determination of average utility at all. Bentham's principal effort, in his discussion of the ethics of mo-

[11] By a process similar to that by which David Lyons, in his *Forms and Limits of Utilitarianism* (New York: Oxford, 1965), has tried to show that rule utilitarianism is equivalent to act utilitarianism.

tives, is to show, by a tiresome profusion of examples, that the application of the test of utility to sorts of motive yields no results, because "there is no sort of motive but may give birth to any sort of action" (*Introduction*, 128); his argument depends on the use of very thin descriptions of sorts of motive.

The doctrine that a type of motive is better, the greater the utility of commending or fostering it in a system of moral education, might seem to be another version of universalistic motive utilitarianism, but is not a form of motive utilitarianism at all. For in it the test of utility is directly applied not to motives or types of motive, but to systems of moral education.

I am not convinced (nor even inclined to believe) that any purely utilitarian theory about the worth of motives is correct. But motive-utilitarian considerations will have some place in any sound theory of the ethics of motives, because utility, or conduciveness to human happiness (or more generally, to the good), is certainly a great advantage in motives (as in other things), even if it is not a morally decisive advantage.

ROBERT MERRIHEW ADAMS

University of California, Los Angeles

BOOK REVIEWS

The Structure of Morality. HECTOR-NERI CASTAÑEDA. Springfield, Ill.: Charles C Thomas, 1974. x, 239 p. $12.75.

This book is to be welcomed for a number of reasons. The first is that most of the very penetrating contributions to moral philosophy of its author, hitherto scattered through the periodicals, are now available for the first time in a convenient form. Another is that, at a time when those who seek to bring philosophy to bear on practical questions are showing an increasing tendency to say "Good bye" to logic and rely on their own and their readers' prejudices, Professor Castañeda draws us back firmly to the truth that a theory of morality, if it is to do what is required of it, has to be securely based on a rigorous study of the logical properties of the moral words and of the other kinds of practical discourse. He says "In this book we break the tradition of attempting to define or analyze morality without bothering to formulate the appropriate underlying logical foundation" (7); and, although he exaggerates the extent to which his own work is innovative in this respect, the