

170
W721m

Morality:

An Introduction to Ethics

Bernard Williams

HARPER TORCHBOOKS V TB 1632

The Amoralist

"Why should I do anything?" Two of the many ways of taking that question are these: as an expression of despair or hopelessness, when it means something like "Give me a reason for doing anything; everything is meaningless"; and as sounding a more defiant note, against morality, when it means something like "Why is there anything that I *should, ought to, do?*"

Even though we can paraphrase the question in the first spirit as "Give me a reason . . .," it is very unclear that we can in fact give the man who asks it a reason—that, starting from so far down, we could *argue* him into caring about something. We might indeed "give him a reason" in the sense of finding something that he is prepared to care about, but that is not inducing him to care by reasoning, and it is very doubtful whether there could be any such thing. What he needs is help, or hope, not reasonings. Of course it is true that if he stays alive he will be doing *something*, rather than something else, and thus in some absolutely minimal sense he has some sort of reason, some minimal preference, for doing those things rather than other things. But to point this out gets us hardly anywhere; he does

those things just mechanically, perhaps, to keep going, and they mean nothing to him. Again, if he sees his state as a reason for suicide, then that would be to make a real decision; as a way out of making any decisions, suicide comes inevitably one decision too late (as Camus points out in *Le Mythe de Sisyphe*). But it would be no victory for us or for him if it turned out there was after all just one decision that he was prepared to acknowledge, that one.

I do not see how it could be regarded as a defeat for reason or rationality that it had no power against this man's state; his state is rather a defeat for humanity. But the man who asks the question in the second spirit has been regarded by many moralists as providing a real challenge to moral reasoning. He, after all, acknowledges some reasons for doing things; he is, moreover, like most of us some of the time. If morality can be got off the ground rationally, then we ought to be able to get it off the ground in an argument against him; while, in his pure form—in which we can call him the *amoralist*—he may not be actually persuaded, it might seem a comfort to morality if there were reasons which, if he were rational, would persuade him.

We might ask first what motivations he does have. He is indifferent to moral considerations, but there are things that he cares about, and he has some real preferences and aims. They might be, presumably, pleasure or power; or they might be something much odder, such as some passion for collecting things. Now these ends in themselves do not exclude some acknowledgment of morality; what do we have to

leave out to represent him as paying no such acknowledgment? Presumably such things as his caring about other people's interests, having any inclination to tell the truth or keep promises if it does not suit him to do so, being disposed to reject courses of action on the ground that they are unfair or dishonorable or selfish. These are some of the substantial materials of morality. We should perhaps also leave out a more formal aspect of morality, namely any disposition on his part to stand back and have the thought that if it is 'all right' for him to act in these ways, it must be 'all right' for others to act similarly against him. For if he is prepared to take this stance, we might be able to take a step towards saying that he was not a man without a morality, but a man with a peculiar one.

However, we need a distinction here. In one way, it is possible for a man to think it 'all right' for everyone to behave self-interestedly, without his having got into any distinctively moral territory of thought at all: if, roughly, 'it's all right' means 'I am not going to moralize about it'. He will be in some moral territory if 'all right' means something like 'permitted', for that would carry implications such as 'people ought not to interfere with other people's pursuit of their own interests', and that is not a thought which, as an amoralist, he can have. Similarly, if he objects (as he no doubt will) to other people treating him as he treats them, this will be perfectly consistent so long as his objecting consists just in such things as his not liking it and fighting back. What he cannot consistently do is *resent* it or

disapprove of it, for these are attitudes within the moral system. It may be difficult to discover whether he has given this hostage to moral argument or not, since he will no doubt have discovered that insincere expressions of resentment and moral hurt serve to discourage some of the more squeamish in his environment from hostile action.

This illustrates, as do many of his activities, the obvious fact that this man is a parasite on the moral system, and he and his satisfactions could not exist as they do unless others operated differently. For, in general, there can be no society without some moral rules, and he needs society; also he takes more particular advantage of moral institutions like promising and of moral dispositions of people around him. He cannot deny, as a fact, his parasitic position; but he is very resistant to suggestions of its relevance. For if we try saying "How would it be for you if everyone behaved like that?" he will reply, "Well, if they did, not good, I suppose—though in fact I might do better in the resulting chaos than some of the others. But the fact is, most of them are not going to do so; and if they do ever get round to it, I shall be dead by then." The appeal to the consequences of an *imagined* universalization is an essentially moral argument, and he is, consistently, not impressed by it.

In maintaining this stance, there are several things he must, in consistency, avoid. One—as we noticed before, in effect—is any tendency to say that the more or less moral majority have *no right* to dislike him, reject him, or treat him as an enemy, if

indeed they are inclined to do so (his power, or charm, or dishonesty may be such that they do not). No thoughts about justification, at least of that sort, are appropriate to him. Again, he must resist, if consistent, a more insidious tendency to think of himself as being in character really rather splendid—in particular, as being by comparison with the craven multitude notably courageous. For in entertaining such thoughts, he will run a constant danger of getting outside the world of his own desires and tastes into the region in which certain dispositions are regarded as excellent for human beings to have, or good to have in society, or such things; and while such thoughts need not lead directly to moral considerations, they give a substantial footing to them, since they immediately invite questions about what is so good about those dispositions, and it will be difficult for him to pursue those questions very far without thinking in terms of the *general* interests and needs of his fellow human beings, which would land him once more back in the world of moral thought from which he is excluding himself.

The temptation to think of himself as courageous is a particularly dangerous one, since it is itself very nearly a moral notion and draws with it a whole chain of distinctively moral reflections. This man's application of the notion will also have a presupposition which is false: namely, that the more moral citizens would be amoral if they could get away with it, or if they were not too frightened, or if they were not passively conditioned by society—if, in general, they did not suffer from inhibitions. It is

the idea that they are afraid that gives him the idea of his own courage. But these presuppositions are absurd. If he means that if as an individual one could be sure of getting away with it, one would break any moral rule (the idea behind the model of Gyges' ring of invisibility in Plato's *Republic*), it is just false of many agents, and there is reason why: the more basic moral rules and conceptions are strongly internalized in upbringing, at a level from which they do not merely evaporate with the departure of policemen or censorious neighbors. This is part of what it is for them to be moral rules, as opposed to *merely* legal requirements or matters of social convention. The effects of moral education can actually be to make people *want* to act, quite often, in a non-self-interested way, and it often succeeds in making it at least quite difficult, for internal reasons, to behave appallingly.

But this, he will say, is just social conditioning; remove that, and you will find no moral motivations. —We can reject the rhetoric of the word 'conditioning'; even if there were a true theory, which there is not, which could explain all moral and similar education in terms of behaviorist learning theory, it would itself have to explain the very evident differences between successful and intelligent upbringing, which produces insight, and the production of conditioned reflexes. But let us say instead that all moral motivation is the product of social influences, teaching, culture, etc. It is no doubt true. But virtually everything else about a man is such a product, including his language, his methods of thought, his tastes, and

even his emotions, including most of the dispositions that the amoralist sets store by. —But, he may say, suppose we grant that anything complex, even my desires, are influenced by culture and environment, and in many cases produced by these; nevertheless there are *basic* impulses, of a self-interested kind, which are at the bottom of it all: these constitute what men are *really* like.

If 'basic' means 'genetically primitive', he may possibly be right: it is a matter of psychological theory. But even if true in this sense, it is once more irrelevant (to his argument, not to questions about how to bring up children); if there is such a thing as what men are *really* like, it is not identical with what very small children are like, since very small children have no language, again, nor many other things which men really have. If the test of what men are *really* like is made, rather, of how men may behave in conditions of great stress, deprivation, or scarcity (the test that Hobbes, in his picture of the state of nature, imposed), one can only ask again, why should that be the test? Apart from the unclarity of its outcome, why is the test even appropriate? Conditions of great stress and deprivation are not the conditions for observing the typical behavior of any animal nor for observing other characteristics of human beings. If someone says that if you want to see what men are *really* like, see them after they have been three weeks in a lifeboat, it is unclear why that is any better a maxim with regard to their motivations than it is with regard to their physical condition.

If there is such a thing as what men are *really* like, it may be that (in these sorts of respects, at least) it is not so different from what they are *actually* like; that is, creatures in whose lives moral considerations play an important, formative, but often insecure role.

The amoralist, then, would probably be advised to avoid most forms of self-congratulatory comparison of himself with the rest of society. The rest may, of course, have some tendency to admire him, or those may who are at such a distance that he does not tread directly on their interests and affections. He should not be too encouraged by this, however, since it is probably wish-fulfillment (which does not mean that they would be like him if they could, since a wish is different from a frustrated desire). Nor will they admire him, still less like him, if he is not recognizably human. And this raises the question, whether we have left him enough to be that.

Does he care for anybody? Is there anybody whose sufferings or distress would affect him? If we say "no" to this, it looks as though we have produced a psychopath. If he is a psychopath, the idea of arguing him into morality is surely idiotic, but the fact that it is idiotic has equally no tendency to undermine the basis of morality or of rationality. The activity of justifying morality must surely get any point it has from the existence of an alternative—there being something to justify it *against*. The amoralist seemed important because he seemed to provide an alternative; his life, after all, seemed to have its attractions. The psychopath is, in a certain

way, important to moral thought; but his importance lies in the fact that he appalls us, and we must seek some deeper understanding of how and why he appalls us. His importance does not lie in his having an appeal as an alternative form of life.

The amoralist we loosely sketched before did seem to have possibly more appeal than this; one might picture him as having some affections, occasionally caring for what happens to somebody else. Some stereotype from a gangster movie might come to mind, of the ruthless and rather glamorous figure who cares about his mother, his child, even his mistress. He is still recognizably amoral, in the sense that no general considerations weigh with him, and he is extremely short on fairness and similar considerations. Although he acts for other people from time to time, it all depends on how he happens to feel. With this man, of course, in actual fact arguments of moral philosophy are not going to work; for one thing, he always has something he would rather do than listen to them. This is not the point (though it is more of a point than some discussions of moral argument would lead one to suppose). The point is rather that he provides a model in terms of which we may glimpse what moral considerations need to get off the ground, even though they are unlikely in practice to get off the ground in a conversation with him.

He gives us, I think, almost enough. For he has the notion of doing something *for* somebody, because that person needs something. He operates with this notion in fact only when he is so inclined; but it is

not itself the notion of his being so inclined. Even if he helps these people because he wants to, or because he likes them, and for no other reason (not that, so far as these particular actions are concerned, he needs to improve on those excellent reasons), what he wants to do is to *help them in their need*, and the thought he has when he likes someone and acts in this way is 'they need help', not the thought 'I like them and they need help'. This is a vital point: this man is capable of thinking in terms of others' interests, and his failure to be a moral agent lies (partly) in the fact that he is only intermittently and capriciously disposed to do so. But there is no bottomless gulf between this state and the basic dispositions of morality. There are people who need help who are not people who at the moment he happens to want to help, or likes; and there are other people who like and want to help other particular people in need. To get him to consider their situation seems rather an extension of his imagination and his understanding, than a discontinuous step onto something quite different, the 'moral plane'. And if we could get him to consider their situation, in the sense of thinking about it and imagining it, he might conceivably start to show some consideration for it: we extend his sympathies. And if we can get him to extend his sympathies to less immediate persons who need help, we might be able to do it for less immediate persons whose interests have been violated, and so get him to have some primitive grasp on notions of fairness. If we can get him all this way, then, while he has no

doubt an extremely shaky hold on moral considerations, he has some hold on them; at any rate, he is not the amoralist we started with.

This model is not meant to sketch the outline of a construction of the whole of morality from the possibility of sympathy and the extensions of sympathy: that would be impossible. (Even Hume, who perhaps came nearest to it, did not attempt that. His system, among the many interesting and valuable things that it contains, has a distinction between the 'natural' and the 'artificial' virtues which is relevant to this point.) The model is meant to suggest just one thing: that if we grant a man with even a minimal concern for others, then we do not have to ascribe to him any fundamentally new kind of thought or experience to include him in the world of morality, but only what is recognizably an extension of what he already has. He is not very far into it, and it is an extensive territory; as we saw in drawing up the amoralist, you have to travel quite a long way to get out of it. But the man with the extended sympathies, the ability to think about the needs of people beyond his own immediate involvement, is recognizably in it.

It does not follow from this that having sympathetic concern for others is a necessary condition of being in the world of morality, that the way sketched is the *only* way 'into morality'. It does not follow from what has so far been said; but it is true.

Some of the considerations touched on here, about moral and other motivations, we shall come back to later. I shall now turn to someone who is also found

upsetting by morality, but in a different way from the amoralist. This man is content that he should have a morality, but points out that other people have different ones—and insists that there is no way of choosing between them. He is the subjectivist.

Subjectivism: First Thoughts

Consider three statements, each of which, in its different way, expresses a view that moral opinions, or moral judgments, or moral outlooks are 'merely subjective':

- (a) "A man's moral judgments merely state (or express) his own attitudes."
- (b) "Moral judgments can't be proved, established, shown to be true as scientific statements can; they are matters of individual opinion."
- (c) "There are no moral facts; there are only the sorts of facts that science or common observation can discover, and the values that men place on those facts."

The three statements come very close to one another, and in discussions of subjectivism and objectivism one often finds versions of the three being used virtually interchangeably. They are, indeed, genuinely related to one another. Yet they are significantly different. The first, (a), expresses what might be called in a broad sense a *logical* or *linguistic* view: it purports to tell us something about what moral remarks are or do. The second, (b), intro-

duces a set of notions not present in the first, notions connected with the concept of *knowledge* and may be taken to express an *epistemological* view about moral judgments. The third statement, (c), is the vaguest and least tangible of the three, and shows on its surface the danger of collapsing, partly or completely, into one or other of the first two: which is what many philosophers would claim it must do.—Yet in its inadequate way it seems to gesture towards something which is closest of all to what has concerned many who have been worried about moral objectivity: the idea that there is no moral order ‘out there’—out there, in the world, are only the sorts of things and the sorts of facts that science, and the more everyday processes of human enquiry of which science is a refinement, deal. *„Alles anderes ist Menschenwerk.* Statement (c) can be said—using the word in an unambitious way—to express a *metaphysical* view.

The metaphysical view brings out most explicitly something that is latent in all the three statements, a distinction between fact and value. A central concern of much modern moral philosophy has been the distinction between fact and value. One important form that this concern has taken has been to emphasize the distinction, while rejecting the supposedly disquieting consequences of it, by trying to show either that they are not consequences, or that they are not disquieting. This project of *defusing* subjectivism (as it might be called) can be expressed in terms of our three subjectivist statements roughly as follows: that, insofar as they are defensible, they

come to the same thing; and that what they come to is both not alarming and essential to the nature of morality (the point that it is essential to the nature of morality seems sometimes to be thought, oddly, to imply just by itself the point that it is not alarming).

This project we shall follow, with interruptions, for some while. It starts as follows. Statement (a), first of all, is either false or harmless. It is false if it claims that moral judgments *state* their utterer's attitudes, in the sense of stating that he has those attitudes. For if this were so, they would be simply autobiographical remarks, replaceable without loss by statements explicitly of the form ‘My attitude towards this is . . .’ or ‘I feel. . .’ But if this were so, there would be no interpersonal moral disagreements; two persons expressing what we would normally take to be conflicting views would not be expressing conflicting views at all, but would be, rather, like two persons on a boat, one of whom says that he feels sick while the other says that he, on the other hand, does not. But it is an evident fact that there are genuine moral disagreements, and that moral views can conflict. Moral judgments must (at least to this extent) mean what we take them to mean; and what we take them to mean, the way we use them, is such that they do not merely make autobiographical claims, but a sort of claim which is being rejected by someone who utters a contrary moral judgment. Thus they do not merely describe the speaker's own attitude.

This argument, however, does not dispose of the

possibility that moral utterances, while they do not directly describe, may nevertheless have the function of expressing, the utterer's attitude. This claim in itself is incontestable and harmless; for in itself it comes to no more than saying that a man who makes a moral judgment can (if that judgment is sincere) be said to be expressing his attitude to a certain moral issue, and this does not support any distinctively subjectivist view of moral judgments: a man who sincerely makes a factual judgment may be said to be expressing his belief about a factual issue, but that does not support a subjectivist view of factual judgments. The subjectivist interpretation comes in when one says that one who makes a moral judgment expresses his attitude, and *that is all there is to be said about it*. In particular the subjectivist force of (a) lies in a suggestion that there is no question of the attitudes expressed in moral judgments being right or wrong, whereas there is a question of the beliefs expressed in factual judgments being true or false.

Now that the thesis has regained a distinctively subjectivist form it will be said that it is false. For—appealing once more to the ways in which moral judgments are actually made and treated—it is not true that there is no question of moral attitudes being right or wrong. One of their distinguishing marks, as against mere expressions of taste or preference, for instance in matters of food, is that we take seriously the idea of a man's being wrong in his moral views; indeed, the very concept of a moral *view* marks a difference here, leaning as it does in

the direction of belief rather than of mere taste or preference. It is precisely a mark of morality that *de gustibus non disputandum* is a maxim which does not apply to it.

To this it might be replied that the fact that moral attitudes can be called 'right' or 'wrong', and that the question of their rightness or wrongness is taken seriously, does not in any ultimate sense help to transcend subjectivism. It shows not that moral attitudes are more than (merely) attitudes but that they are attitudes which we get disturbed about; that it matters to us to secure similarity of attitude within society. The use of the language of 'right' and 'wrong' can be seen as part of the apparatus of securing agreement, marking off those who disagree and so forth; it remains the case that all we have are people's attitudes toward different sorts of conduct, personality, social institutions, etc.

Yet this account, in its turn, seems inadequate to the facts: it must at least underestimate the logical complexity of the situation. Notably it fails to account for the undoubted facts that a man may be in a state of moral doubt, which he may resolve—that a man can nonarbitrarily change his mind about a moral matter, not merely in the individual case, but on a general issue, and for reasons. Thus a man previously convinced that a permissive attitude towards abortion was wrong, might change his view on this, and not merely (for instance) because he felt lonely in a group which held the permissive attitude. No doubt many writers on moral philosophy overestimate the extent to which people are led by

rational considerations to modify their moral views; those writers ignore the evident extent to which attitudes are modified by factors such as the desire to conform with one group or nonconform with another—the groups themselves not being chosen in the light of moral reasons, but rather determined by the individual's situation and needs. But this is, in one sense, beside the point. For even if moral attitudes were rarely *determined* by reasons, and the reasons advanced in their support were rationalizations, our model of moral attitudes and moral judgments must at least be complex enough to leave a place for the rationalizations. It is only if the position to which a man is led by these forces satisfies some conditions of being the sort of position to which reasons are relevant that we can understand it as a moral position at all.

However, even if moral attitudes are of a character to admit of support and attack by reasons, and the deployment of reasons in reaching a conclusion; nevertheless it may be said that these activities are possible only within a framework of given assumptions. If people can argue one with another about an individual moral issue or a question of principle, this is only because there are moral attitudes in the background which are not in dispute and in the light of which the argument goes on. Putting this point in a rather stronger form, it might be said that it is only about the *application* of accepted moral views that the argument goes on. So where there is no background of moral agreement, there can be no argument. At this point the subjectivist attitude can

reappear, claiming now that all that has been shown by the considerations about exchanging reasons is that the morality of a man or a society is to some degree general and systematic and that general attitudes can be applied to less general cases. When we get outside the framework of agreed general attitudes, there is no further argument, and no way of showing any position to be right or wrong.

This comes very close to our second formulation, (b), of subjectivism; we have been led to it, it seems, by modifications of the first formulation. One might, however, want to change one element that came naturally into the formulation of (b), and, with that, one implication of the term 'subjectivism'. For when one turns to the issue of ultimate disagreements, it is natural to take as the unit which holds a set of moral attitudes, the society rather than the individual—not in order to hypostatize societies, but to draw attention to the point that there are limits to the degree of ultimate disagreement that can exist within a society (for without some degree of moral homogeneity it would not be a society); but there are no limits, at least of that kind, on disagreement between societies.

Interlude: Relativism

Let us at this stage of the argument about subjectivism take a brief rest and look round a special view or assemblage of views which has been built on the site of moral disagreements between societies. This is *relativism*, the anthropologists' heresy, possibly the most absurd view to have been advanced even in moral philosophy. In its vulgar and unregenerate form (which I shall consider, since it is both the most distinctive and the most influential form) it consists of three propositions: that 'right' means (can only be coherently understood as meaning) 'right for a given society'; that 'right for a given society' is to be understood in a functionalist sense; and that (therefore) it is wrong for people in one society to condemn, interfere with, etc., the values of another society. A view with a long history, it was popular with some liberal colonialists, notably British administrators in places (such as West Africa) in which white men held no land. In that historical role, it may have had, like some other muddled doctrines, a beneficent influence, though modern African nationalism may well deplore its tribalist and conservative implications.

Whatever its results, the view is clearly inconsistent, since it makes a claim in its third proposition, about what is right and wrong in one's dealings with other societies, which uses a *nonrelative* sense of 'right' not allowed for in the first proposition. The claim that human sacrifice, for instance, was 'right for' the Ashanti comes to be taken as saying that human sacrifice was right among the Ashanti, and this in turn as saying that human sacrifice among the Ashanti was right; i.e., we have no business to interfere with it. But this last is certainly not the sort of claim allowed by the theory. The most the theory can allow is the claim that it is right for (i.e., functionally valuable for) our society not to interfere with Ashanti society, and, first, this is certainly not all that was meant, and, second, is very dubiously true.

Apart from its logically unhappy attachment of a nonrelative morality of toleration or noninterference to a view of morality as relative, the theory suffers in its functionalist aspects from some notorious weaknesses of functionalism in general, notably difficulties that surround the identification of 'a society'. If 'society' is regarded as a cultural unit, identified in part through its values, then many of the functionalist propositions will cease to be empirical propositions and become bare tautologies: it is tediously a necessary condition of the survival of a group-with-certain-values that the group should retain those values. At the other extreme, the survival of a society could be understood as the survival of

certain persons and their having descendants, in which case many functionalist propositions about the necessity of cultural survival will be false. When in Great Britain some Welsh nationalists speak of the survival of the Welsh language as a condition of the survival of Welsh society, they manage sometimes to convey an impression that it is a condition of the survival of Welsh people, as though the forgetting of Welsh were literally lethal.

In between these two extremes is the genuinely interesting territory, a province of informative social science, where there is room for such claims as that a given practice or belief is integrally connected with much more of a society's fabric than may appear on the surface, that it is not an excrescence, so that discouragement or modification of this may lead to much larger social change than might have been expected; or, again, that a certain set of values or institutions may be such that if they are lost, or seriously changed, the people in the society, while they may physically survive, will do so only in a deracinated and hopeless condition. Such propositions, if established, would of course be of first importance in deciding what to do; but they cannot take over the work of deciding what to do.

Here, and throughout the questions of conflict of values between societies, we need (and rarely get) some mildly realistic picture of what decisions might be being made by whom, of situations to which the considerations might be practically relevant. Of various paradigms that come to mind, one is that of

conflict, such as the confrontation of other societies with Nazi Germany. Another is that of control, where (to eliminate further complications of the most obvious case, colonialism) one might take such a case as that of the relations of the central government of Ghana to residual elements of traditional Ashanti society. In neither case would functionalist propositions in themselves provide any answers at all. Still less will they where a major issue is whether a given group should be realistically or desirably regarded as 'a society' in a relevant sense, or whether its values and its future are to be integrally related to those of a larger group—as with the case of blacks in the United States.

The central confusion of relativism is to try to conjure out of the fact that societies have differing attitudes and values an *a priori* nonrelative principle to determine the attitude of one society to another; this is impossible. If we are going to say that there are ultimate moral disagreements between societies, we must include, in the matters they can disagree about, their attitudes to other moral outlooks. It is also true, however, that there are inherent features of morality that tend to make it difficult to regard a morality as applying only to a group. The element of universalization which is present in any morality, but which applies under tribal morality perhaps only to members of the tribe, progressively comes to range over persons as such. Less formally, it is essential (as was remarked earlier) to morality and its role in any society that

certain sorts of reactions and motivations should be strongly internalized, and these cannot merely evaporate because one is confronted with human beings in another society. Just as *de gustibus non disputandum* is not a maxim which applies to morality, neither is 'when in Rome do as the Romans do', which is at best a principle of etiquette.

Nor is it just a case of doing as the Romans do, but of putting up with it. Here it would be a platitude to point out that of course someone who gains wider experience of the world may rightly come to regard some moral reaction of his to unfamiliar conduct as parochial and will seek to modify or discount it. There are many important distinctions to be made here between the kinds of thoughts appropriate to such a process in different cases: sometimes he may cease to regard a certain issue as a moral matter at all, sometimes he may come to see that what abroad looked the same as something he would have deplored at home was actually, in morally relevant respects, a very different thing. (Perhaps—though one can scarcely believe it—there were some missionaries or others who saw the men in a polygamous society in the light of seedy bigamists at home.) But it would be a particular moral view, and one both psychologically and morally implausible, to insist that these adaptive reactions were the only correct ones, that confronted with practices which are found and felt as inhuman, for instance, there is an *a priori* demand of acceptance. In the fascinating book by Bernal de Diaz, who went with Cortez to Mexico, there is an account of what

they all felt when they came upon the sacrificial temples. This morally unpretentious collection of bravos was genuinely horrified by the Aztec practices. It would surely be absurd to regard this reaction as merely parochial or self-righteous. It rather indicated something which their conduct did not indicate, that they regarded the Indians as men rather than as wild animals.

It is fair to press this sort of case, and in general the cases of actual confrontation. 'Every society has its own standards' may be, even if confused, a sometimes useful maxim of social study; as a maxim of social study it is also painless. But what, after all, is one supposed to do if confronted with a human sacrifice?—not a real question for many of us, perhaps, but a real question for Cortez. "It wasn't their business," it may be said; "they had no right to be there anyway." Perhaps—though this, once more, is necessarily a nonrelative moral judgment itself. But even if they had no right to be there, it is a matter for real moral argument what would *follow* from that. For if a burglar comes across the owner of the house trying to murder somebody, is he morally obliged not to interfere because he is trespassing?

None of this is to deny the obvious facts that many have interfered with other societies when they should not have done; have interfered without understanding; and have interfered often with a brutality greater than that of anything they were trying to stop. I am saying only that it cannot be a consequence of the nature of morality itself that no society ought ever to interfere with another, or that

individuals from one society confronted with the practices of another ought, if rational, to react with acceptance. To draw these consequences is the characteristic (and inconsistent) step of vulgar relativism.

Subjectivism: Further Thoughts

The enticements of vulgar relativism are not very tempting. But its central error is quite important; and that it is an error has a significance beyond that particular doctrine, concerned as that is with the relations between societies. It is possible for someone persuaded of subjectivist views to cease to care about moral issues. (This is different from ceasing to regard something as, in itself, a moral issue. Thus it is both possible and reasonable to suppose that there is no distinctively sexual morality, in the sense of moral considerations that govern sexual relationships and nothing else; while admitting the extremely obvious fact that sexual relationships are profoundly and specially liable to give rise to moral issues, of trust, exploitation, unconcern for the interests of third parties and so forth.) A man confronted with some monstrous political injustice, for instance, may feel no confidence in protesting or fighting against it because, as he says, "Who's to judge?" or "It's only my feelings against theirs," or something like that.

Insofar as there is any traceable intellectual link between the subjectivism and the indifference, it must involve something like the relativist error. For

the thought must be something like this: "Because subjectivism is true, I am not justified in protesting." If this is right, it must be so either because, if subjectivism is true, no one is justified in doing anything; or because, if subjectivism is true, he is specially not justified in protesting. If the first of these, then the inflictors of the injustice are not justified in what they are doing either, nor is this man justified in *not* protesting, and these considerations remove any *basis* he was supposed to have for his indifference. In any case, the argument in this version has surely gone too far, even for subjectivism; since subjectivism did not claim it to be impossible to think that anything was justified, but only that when a given man thinks something is justified, he cannot in the end be proved wrong. The man we are arguing with has somehow got from that to a position in which no one (however subjectively) can think something justified, and that must be a mistake; unless, of course, subjectivism is inconsistent, in which case subjectivism is false, and the argument is over anyway.

Let us try the second alternative, then, that because subjectivism is true, he is specially not justified in protesting. Why should this be so? "Well," he may say, "they think they are right, and who am I to say that they are wrong?" But the apparent force of this is entirely gained from its subtly moving out of the subjectivist arena and importing the idea that there is such a thing as objective rightness, only he is not sure whether these other people's actions possess it or not. Sticking to the subjectivist path, he

must recognize that if he chooses to think that they are wrong and that he is right in protesting, then no one can say he is wrong either, and he can be no less justified in protesting than they are in doing what they are doing. Another way of putting this point is this: "perhaps they are right" must be one of *his* moral thoughts. If he also has the thought "They are wrong (only I am not justified in protesting)," he has inconsistent moral thoughts within his own system, and subjectivism never required him to do that. In fact, of course, the thought "Who am I to say that they are wrong?" is one that is had, not within his own subjective compound, but in midair above his own and the other people's; it tries to stand outside all moral positions (including the thinker's own) and yet still be a moral thought. But this midair place, by subjectivism itself, is not a place in which anyone can have a moral thought.

These sorts of reasons show why the defusers of subjectivism say that it leaves everything where it was and cannot possibly logically imply indifferentism or any other practical attitude. (They move perhaps rather rapidly from that to the claim that their views cannot *encourage* any such attitudes, perhaps on the questionable ground that I cannot be responsible for muddles that people get into as a result of what I say.) But does subjectivism leave everything where it was? Surely not *everything*. The dialectical skeins we just went through are, I think, perfectly valid, as showing that indifferentism could not follow from subjectivism. But the reason why this is so is that the argument for indifferentism

requires, and subjectivism forbids, the midair position. But in forbidding the midair position, subjectivism *seems* to have taken something away, for we at least seem to recognize the midair position (for instance, in subjectivism's own statements). Another way of putting our bafflement is that we seem in these arguments to have been given no special reason why the midair position is debarred to *moral-ity*. It was said that the midair position was no place for a moral thought. But it does seem to be a place for some sorts of thought, in particular, factual thought. Indeed, subjectivism itself (see (b) and (c) of our original formulations, given above), insists on the contrast between moral attitudes and factual beliefs, regarding the latter as 'objective' in a way that the first are not. The subjectivist leaves us, and may mean to leave us, with an uneasy feeling that factual beliefs have got, and moral beliefs lack, something that it is nice to have: that factual beliefs and science are somehow *solidier* than morality.

But still, it will be said, subjectivism leaves everything where it was, so far as morality is concerned, though not perhaps so far as muddled feelings about morality are concerned. Even granted the contrast of solidity we cannot draw any practical conclusions. We cannot, in particular, conclude (as some today are obviously inclined to conclude) that since science is objective and morality is not, we are objectively justified in *devoting ourselves* to science, while only subjectively justified in protesting against injustice. For devoting oneself to science is as much a practical activity as any other, and there is no more rea-

son why that one should be objectively justified rather than any other. Justifications for doing objective subjects are not objective justifications for doing those subjects, any more than the fact that there are deductive justifications of the theorems of *Principia Mathematica* means that there are deductive justifications of the projects of reading, rehearsing, or discovering the theorems of *Principia Mathematica*. All these are instances of that strangely tempting fallacy, the 'fat oxen' principle: who drives fat oxen must himself be fat.

Equally, though more subtly, the fact that 'pragmatic' political policies (i.e., policies which apply sophisticated technical considerations to the pursuit of self-interest) can be evaluated one against another with greater expectations of expert agreement than more idealistic policies often achieve, does not, even if true, show that we are on more objective and solid ground in adopting pragmatic rather than idealistic policies. Once we have adopted them we may be on more objective ground in working them, i.e., in deciding what to do next; but we never get on to more objective ground with reference to the question whether we ought to be doing this sort of thing at all. We are merely more comfortable, and if among the sophisticated experts, in work.

So subjectivism, even with its unsettling contrast between morality and science, is still not logically committed to producing practical consequences. It cannot even yield the conclusion that we are more solidly justified in *having* factual beliefs than moral ones; or on more objective ground in pursuing a

factual question than pursuing a moral one; or objectively justified in seeking the truth about anything; or in trying to find scientific explanations of phenomena rather than resting content with superstition. For all it gave us was that factual and scientific beliefs were objective; that we should seek factual or scientific beliefs is not itself a factual or scientific belief.

Perhaps the subjectivist will readily grant these points about engaging in scientific activity or rational factual thought. Let us then look more closely at his central contrast between those activities on the one hand, and moral thought on the other. He may well start by saying that if we engage in such activities, then, facts being as they are, we are bound to reach some agreed scientific or factual beliefs; but we are not bound, because we engage in moral thought, and the facts are so, to reach some agreed moral beliefs. There is one element in this answer which has to be corrected straightaway. For it is of course not true that if the facts are so, we are bound (granted factual thought) to reach some agreed factual beliefs: the facts may be hidden from us. The most we can say is that if we recognize that the facts are so, we are bound to reach some agreed factual beliefs. And this is a tautology, since our recognizing the facts to be so entails our reaching some agreed factual beliefs. Perhaps the subjectivist can improve on this unexciting proposition by saying something like this: that if two observers are in the same observational situation, and have the same concepts and are not defective as observers, etc., then they will

reach the same factual beliefs about that situation. If the "etc." in this can be handled so as to make the statement true, it will almost certainly emerge as necessarily true as well. All right, says the subjectivist; but *this* is not necessarily true, indeed is not true at all: that if two observers are in the same observational situation, have the same concepts, etc., then they are bound to reach the same *moral* beliefs—and there is the contrast. But, we may say, are we sure that this latter is not true if we posit that they have the same *moral concepts*—which is surely the fair parallel? No, it is not true even then, the subjectivist says; for if they have, as they might, rather minimal moral concepts, such as merely the concept of *what one ought to do*, they can agree on all the facts and disagree morally.

This is a central position of many philosophers who insist on the fact/value distinction. They represent different moral outlooks as all using some common, skeletal, moral concept and giving it different fillings or contents. Now we might point out that a lot of moral thought does not operate with such impoverished concepts, that we standardly think in terms of more complex ideas of virtues, types of wrong action, etc., notions of *thrift*, for instance, or *cowardice*, or *loyalty*, or *the duties of one's job*. With those more substantial concepts, there is much greater hope that if we use the same concepts, we will reach agreement, or at least the sorts of disagreement which we may reasonably expect, and which are less unnerving. But the subjectivist will say that there is this asymmetry, that in scientific or

factual contexts, if two observers have different concepts, they may ultimately find either that the concepts are in fact equivalent, or else find reason for preferring one to the other in terms of predictive success, explanatory power, and so forth, and there is no parallel to this in the moral case.

Even if we abandon the naïve view (which some subjectivists, in working their contrast, are attached to) that science 'proves' things; even if we accept that what science does is eliminate hypotheses and that there are infinitely many hypotheses which have never been eliminated because they are too dotty for anyone to bother to test them (a point I have heard Hilary Putnam make); nevertheless, there are eliminatory procedures which must be respected by persons within the scientific framework, and even the conceptual disagreements may with luck yield to impersonally accepted criteria. Yet for some disagreements between persons or groups both engaged in recognizably moral thought, no such procedures exist. The contrast—some contrast—exists.

But why should it not? This, finally, might be the point at which a philosopher who had been arguing with the subjectivist all this time might at last turn round and say: of course the contrast exists; morality is not just like science or factual knowledge, and it is essential that it should not be. The point of morality is not to mirror the world, but to change it; it is concerned with such things as principles of action, choice, responsibility. The fact that men of equal intelligence, factual knowledge, and so forth, confronted with the same situation, may morally

disagree shows something about morality—that (roughly) you cannot pass the moral buck on to how the world is. But that does not show (as subjectivism originally seemed to insinuate) that there is something *wrong* with it.

Some such statement—and I have only blocked it in very roughly—will express the culmination of the project of what I called 'defusing' subjectivism. It has, I think, to be granted some success. The most obvious ways in which somebody might be unnerved by subjectivism seem to have been blocked. Thus, to revert once again to the indifference we discussed before, the 'defuser' can arrange some of the subjectivist's materials into the following argument. We observe that when men of equal scientific or historical competence, equal perceptual and intellectual abilities, etc., disagree strongly about some scientific or historical matter, there is good reason for them to stop disagreeing so strongly, and recognize something which their very disagreement, granted their knowledge and skills, reveals, namely that the matter is *uncertain*: it is rational for them, and third parties, to suspend judgment. One may be tempted to think that the same should obtain in cases of moral disagreement; but this will be a mistake. It depends on first contrasting morality and factual knowledge, and then assimilating them. For the vital difference is that the disagreement in morality involves what should be done, and involves, on each side, caring about what happens; and once you see this difference, you see equally that it could not possibly be a requirement of rationality that you

should stop caring about these things because someone else disagrees with you.

This, and the similar arguments, do seem to me to show that the defusing operation, in certain vital respects, has succeeded. Has it totally succeeded? If it has, then we were wrong, some time back, in feeling uneasy because subjectivism banned, for morality but not for factual beliefs, something we called the 'midair position'. But, I think, we were not entirely wrong, in feeling that unease—perhaps we can now see why.

If I have a factual disagreement with a man, I might think: "I believe that *p*, he believes that not-*p*. Perhaps he is right." The natural way of taking this is as an expression of doubt, of somewhat shaken confidence; the way I mean "perhaps he is right" implies that while I still believe that *p*, I do not believe it as strongly as all that. In this way of taking it, it can be painlessly paralleled for the moral case; for if I think "perhaps he is right," this will naturally be taken as an expression of somewhat shaken confidence also. But now, in the factual case, there is a possible thought which seems rather similar, but is not the same: the thought "I am convinced that *p*, but it is possible for all that that not-*p*," where this is not an expression of *doubt* at all, but rather registers the impersonal consideration that how things are is independent of my belief; however they are, they are, whatever I believe. We do not know exactly what content to assign to this thought, but unless we have the most drastic philosophical views, we are convinced that it has a content: and the defuser will

agree. But even defused subjectivism leaves no parallel thought possible on the moral side: for subjectivism, however defused, there just is no content to "I am convinced that racial discrimination is intrinsically wrong, but it is possible for all that, that it is not," except things like "How convinced am I?" or "I suppose somebody might make me change my mind."

Such a contrast (and it, and related contrasts, need proper investigation, which we cannot attempt here) might make us agree with the third subjectivist formula we originally introduced: there are no ethical facts. Yet once more the defuser will say: this is just one more formulation of what I said already, and of something essential, not detrimental, to morality. For I have already said that moral thought is essentially practical; it is not its business to mirror the world. —But now we might reply: you said that it was not its business to mirror the world of empirical facts, and we agreed. But did we agree that it mirrored no facts at all? And here the locus of our dissatisfaction may become clearer in the thought that the reason why even defused subjectivism seems to have left something out is that moral thinking *feels* as though it mirrored something, as though it were constrained to follow, rather than be freely creative. When we see further that many defusing philosophers express the essential difference between factual and moral thought in terms of a contrast between the intellect and the *will*, and represent the responsibilities of morality in terms of our *deciding* on certain moral principles—then we

have reason to be dissatisfied, either with them, or, if they are right, with moral thought. For certainly the consciousness of a principle of action as freely decided upon is very unlike the consciousness of a moral principle, which is rather of something that has to be *acknowledged*. If it is then said that there is just a psychological explanation of that—then moral thought seems a cheat, presenting itself to us as too like something which it is not.

These remarks only gesture towards a center of dissatisfaction. They leave almost everything to be done: and not perhaps first in moral philosophy. For instance, we need to distinguish two things run together in what was just said, the idea of realism—that thought has a subject matter which is independent of the thought—and the idea of thought being constrained to certain conclusions. Thus mathematical thought has the latter property, but it is a deep and unsettled question in the philosophy of mathematics how far we should or can think in realist terms about the subject matter of mathematics.

Here I shall leave the direct discussion of subjectivism and issues that it raises, with the conclusion that defused subjectivism does not leave everything where it was, but that it does leave more where it was than we might have thought when we started. If subjectivism, however defused, is there, things are with morality not quite as they seemed; but the fraud, we might say, justifies at most resentment rather than panic. We shall not however lose sight of

the idea of constraints on moral thought, the limitations on the creation of values. We shall come round to them by another route, through the idea of goodness. To talk about goodness, we shall start with 'good'.

sized this as opposed to moralities which want as much as possible tidied up, might still be said to be concerned with men's well-being. Something will still be excluded by the use of this term: systems of values or precepts which paid no attention at all to what we can understand men as needing or wanting.

Utilitarianism

In discussing the question whether any moral outlook must ultimately be concerned with human happiness, I have not supposed that question to be the same as the question whether all moral outlooks must be one or another version of *utilitarianism*. Obviously they are not the same question if we take the narrowest sense of 'utilitarianism', which holds that there is just one moral principle, to seek the greatest happiness of the greatest number; that 'happiness' here means pleasure and the absence of pain; and that the one moral principle—since it is the one moral principle—is to be applied to each individual situation ('act-utilitarianism'). Obviously there are all sorts of ways in which a morality can be ultimately concerned with human happiness without being identical with *that*. But I think also that there are all sorts of ways in which morality can be ultimately concerned with human happiness without being identical with utilitarianism even taken in a more extended sense.

A difficulty in discussing this issue is a lack of agreement about how extensively the term 'utilitarianism' may properly or sensibly be used. The term has sometimes been used to include moral outlooks

which do not have anything specially to do with happiness or pleasure at all; in this sense, it is used to refer to any outlook which holds that the rightness or wrongness of an action always depends on the consequences of the action, on its tendency to lead to intrinsically good or bad states of affairs. This very broad sense—which is probably better represented by the word ‘consequentialism’ than by ‘utilitarianism’—is not my concern here; we are interested only in views of this kind which do take *happiness* as the one intrinsically good thing, at which actions and social arrangements are supposedly aimed. But that restriction still leaves a lot of room for different sorts of utilitarianism.

To discuss in a vacuum what might or might not count as a recognizable form of utilitarianism would be a purely verbal and pointless exercise. The question can only be approached by asking what the *point* of the utilitarian outlook on morality is; and that can be discovered not merely, nor principally, by consulting what Bentham and J. S. Mill and other classical exponents of the system had in mind, but by considering what the attractions of the utilitarian outlook are for moral thought. I think that there are four major ones: this is not to deny that these are, in ways worth exploring, related to one another. First, it is non-transcendental, and makes no appeal outside human life, in particular not to religious considerations. It thus helps, in particular, with the entirely reasonable demand that morality now should be obviously free from Christianity. It can even seem to help—because of a certain conservatism which I shall

consider later—with a demand far less reasonable, indeed rightly perceived by Nietzsche to be idiotic, that the morality thus freed from Christianity should be very much the same as the one previously attached to Christianity. In more radical hands, however, utilitarianism promises more radical change.

Second, its basic good, happiness, seems minimally problematical: however much people differ, surely they at least all want to be happy, and aiming at as much happiness as possible must surely, whatever else gives way, be a reasonable aim. Now there is a notorious problem at this point about the transition from a supposedly indisputable aim of seeking one's own happiness, to a more disputable aim of seeking other people's happiness, and the unfortunate Mill has been repeatedly beaten over the head by critics for (it is said) trying to make this transition by deductive argument. I doubt whether that was what he was trying to do, but in any case the problem is of no special force against utilitarianism—there is no reason why it, any more than anyone else, should possess a magic formula for arguing the amoralist out of his amorality. The point is rather that utilitarianism is a *minimum commitment* morality, in this as in other respects: given merely the minimum requirements for being in the moral world, a willingness to consider other people's wants as well as one's own, utilitarianism can get going on this spot. A much more interesting question is whether the ‘indisputable’ aim of happiness can in fact be made to serve utilitarian purposes. We have already seen some reason, in the previous section, for doubting

whether happiness *must* be seen as the aim of human life at all; but even waiving those questions, it is far from clear that any sense in which it is (more or less) indisputably such an end, is also a sense in which utilitarianism can be made to work on it. This is a central issue: we shall be in a better position to consider it when we have looked at the third and fourth attractions of utilitarianism.

Its third attraction is that moral issues can, in principle, be determined by empirical calculation, of consequences. Moral thought becomes empirical, and on questions of public policy, a matter of social science. This has always been found by many one of the most gratifying features of utilitarianism. It is not that the calculations are thought to be easy, or even practically possible in many cases; the charm lies rather in this, that the nature of the difficulty is at least quite unmysterious. All moral obscurity becomes a matter of technical limitations.

Fourth, utilitarianism provides a common currency of moral thought: the different concerns of different parties, and the different sorts of claims acting on one party, can all be cashed (in principle) in terms of happiness. This provision, importantly, has the consequence that a certain kind of conflict, well-known to some other moral outlooks, is impossible—the conflict, that is to say, of two claims which are both valid and irreconcilable. Under some other systems, a man may come to be in a situation in which (as it seems to him) whatever he does involves doing something wrong. For utilitarianism, this is impossible. The various claims he may feel on him can be

brought to the common measure of the Greatest Happiness Principle, and there can be no coherent idea of a right or wrong thing to do, other than what is, or is not, *the best thing to do on the whole*: and if two courses come out equal, then it really cannot matter which he does. As against this, many people can recognize the thought that a certain course of action is, indeed, the best thing to do on the whole in the circumstances, but that doing it involves doing something wrong. This is a thought which for utilitarianism must, I think, ultimately be incoherent. This is one reason for saying (what is certainly true) that for utilitarianism, tragedy is impossible; but it has wider, if not deeper, consequences than that.

The utilitarian may be able to move back a little towards this type of thought, by invoking such things as the desirable social consequences of people being a bit squeamish about certain actions, even when those are in the circumstances, the best available: we shall come back to that type of argument later on. But what he is bound to do as a utilitarian is to regard as an indisputable general aim of moral thought, the reduction of conflict, the elimination wherever possible of value conflicts insoluble without remainder. Here, as elsewhere, he is concerned with efficiency: the generation of conflict is a sign of inefficiency in a value system, and utilitarianism has some might wonder whether such efficiency was an indisputable aim. One can certainly reduce conflict, and make life simpler, by cutting down the range of claims one is prepared to consider; but in certain cases, that might seem not so much a triumph for

rationality, as a cowardly evasion, a refusal to see what is there to be seen (we may ask here, once more, whether defused subjectivism really leaves everything where it was).

So even the attractiveness of utilitarianism's fourth attraction may be importantly disputed. Other difficulties crowd in when one considers what it presupposes. For we are going to be able to use the Greatest Happiness Principle as the common measure of all and everybody's claims, only if the 'happiness' involved is in some sense *comparable* and in some sense *additive*. Only if we can compare the happinesses involved for different people and over different outcomes, and also put them together into some kind of General Happiness, can we make the thing work. At a technical level these problems have been the concern of such subjects as welfare economics and preference theory, which have labored within very artificial assumptions and with only moderate success to deal with them for economic theory. Here we are concerned with more general difficulties. If the 'happiness' involved is to be such as to allow utilitarianism to deliver on its third and fourth promises, can it also be the indisputable aim which was promised in the second?

The answer to that seems to be just 'no'. Bentham offered an account of happiness, namely as pleasure and the absence of pain, which was supposed very clearly to deliver on all the promises at once; but even if it had satisfied (as of course it did not) the conditions of being calculable, comparable, and additive, it failed the condition of being an indisputable

objective: the more it looked like the sort of pleasure that could conceivably be dealt with in those quasi-arithmetical terms, the less it looked like something that any rational man must evidently be aiming at—as Mill came, if uneasily, to see. If, on the other hand, the conception of happiness is made generous enough to include anything that might reasonably be aimed at as a satisfying life or ingredient of a life—then it less and less looks like something which could fit in with the third and fourth conditions. Apart from anything else, there is the difficulty that many things which people actually include in the content of a happy life are things which essentially involve other values, such as integrity, for instance, or spontaneity, or freedom, or love, or artistic self-expression; and not only can they not be treated in the way that the third and fourth conditions require of utilitarianism's 'happiness', but there seems, in the case of some of them at least, an actual contradiction in thinking of them as something that could be so treated.

This is the first general difficulty, then, with utilitarianism. Its 'happiness' has to satisfy certain conditions, if the point of utilitarianism is to be retained; and the condition, that it should be indisputably an aim of human aspiration, conflicts with the other conditions which it must satisfy if it is to be treated as utilitarianism requires it to be treated. Faced with this general difficulty, one way in which utilitarianism tends to react is to dispute the values involved in the more intractable conceptions of happiness, as irrational, perhaps, or as hangovers of a

past age. Such arguments may involve some interesting points on the way, but their strategy is shamelessly circular: utilitarian rationality is made the test of what counts as happiness, in order to remove that sort of happiness which constitutes an objection to utilitarianism. All that is needed to counter this at the theoretical level is a suitable unwillingness to be bullied.

The problem, however, is not confined to the theoretical level: it occurs drastically at the social level, and an unwillingness to be bullied may here be inadequate, or hard to enforce. In cases of planning, conservation, welfare, and social decisions of all kinds, a set of values which are, at least notionally, quantified in terms of resources, are confronted by values which are not quantifiable in terms of resources: such as, the value of preserving an ancient part of a town, or of contriving dignity as well as comfort for patients in a geriatric unit. Again and again defenders of such values are faced with the dilemma, of either refusing to quantify the value in question, in which case it disappears from the sum altogether, or else of trying to attach some quantity to it, in which case they misrepresent what they are about and also usually lose the argument, since the quantified value is not enough to tip the scale. In such matters, it is not that utilitarians are committed to thinking that these other values do not matter; nor are they confined to thinking valuable those things which can presently be handled by cost-benefit analysis. They are perhaps not even bound to think that every social value should eventually be handleable by

something like cost-benefit analysis: they might say that they were not committed to the view that the common currency of happiness is money. But they are committed to something which in practice has those implications: that there are no ultimately incommensurable values. Nor is it an accidental feature of the utilitarian outlook that the presumption is in favor of the monetarily quantifiable, and the other values are forced into the apologetic dilemma we have just met. It is not an accident, because (for one thing) utilitarianism is unsurprisingly the value system for a society in which economic values are supreme; and also, at the theoretical level, because quantification in money is the only obvious form of what utilitarianism insists upon, the commensurability of value.

There is great pressure for research into techniques to make larger ranges of social value commensurable. Some of the effort should rather be devoted to learning—or learning again, perhaps—how to think intelligently about conflicts of values which are incommensurable.

These have been difficulties in satisfying utilitarianism's conditions for 'happiness'. The fact that there are these difficulties does not mean, of course, that we can never in any particular case or sort of case arrive at an idea of what the utilitarian solution would be, or of what sorts of thing the utilitarian would count. Something would be wrong if this did follow, since clearly we can sometimes do these things: in discussing utilitarianism, we are discussing *something*, and something quite often recogniz-

able. So let us grant that in some cases, at least, we do know what is meant by working out what course of action would lead to the greatest happiness all round. Granted this, we now meet two new difficulties. One is, that the process of working out such consequences is itself an activity, which itself in various circumstances possesses various degrees of utility; and this has to go into the sum. The other is, that the answer reached by utilitarian calculation of the particular case seems in certain cases to be morally the wrong answer. There exists a kind of utilitarianism, called Rule-Utilitarianism, whose aim is precisely to solve both these difficulties at once, by one and the same device.

The first problem is, that any actual utilitarian calculation will take place under conditions of considerable uncertainty and very partial information, so that its results are likely to be unreliable. Moreover, the business of calculation itself takes time; and the disposition to calculate in each case has psychological features which may as a matter of fact impede things which are utilitarianly desirable, such as resolute action. These things being so, it is suggested that better consequences may follow from the practice, not of agents calculating each action, but of their subscribing to certain rules which they apply usually without calculation to particular cases; it is the adoption of these rules which is assessed by appeal to the Greatest Happiness Principle, and not the choice of particular actions.

The same idea is invoked to explain the other fact which otherwise presents a difficulty, that we can

easily construct cases—for instance where the conviction of an innocent man is necessary and sufficient to avoid great harms—in which the utilitarian result seems to conflict with what many would regard as the morally right answer: as in this case with justice, so sent difficulties under the act-utilitarian interpretation. The rule-utilitarian can, it is hoped, dissolve these difficulties by claiming that all that has to be shown is that the rules or practices of justice, promise-keeping or truth-telling possess positive utility over the alternatives.

This is only a sketch. Many importantly different things can be covered by the term 'rule-utilitarianism', and different things need to be said about them.⁵ All I will try to do here is to suggest one or two points about how far the utilitarian can consistently go in the rule direction; and argue that either he cannot go far enough to solve the second difficulty, or else he has to go so far that he (and everyone else) ceases to be a utilitarian.

It is certainly possible for a utilitarian, without inconsistency, to adopt a general practice for dealing with a certain kind of case, even though some particular applications of the practice produce a result different from what would have been reached by individual calculation of those instances. The paradigm of this is the accounting system of many public utilities, who may occasionally send out a bill for one dol-

5. For a detailed and subtle discussion of the issues involved, see David Lyons's important book, *The Forms and Limits of Utilitarianism* (Oxford: Clarendon Press, 1969).

lar, even though each bill takes two dollars to process: the point being, that it is actually cheaper to send out all bills when due, however small the amount, rather than to interrupt the processes to extract a few bills. Let us call this the 'gas bill model'.

Now the gas bill model deals in actual consequences: the actual consequences of merely applying a rule, on the one hand, and of making a particular interference, on the other. This sort of model, at least, cannot render palatable to a consistent utilitarian a form of argument which invokes neither the actual consequences of a particular choice, nor the actual consequences of the general following of a rule, but the *hypothetical* consequences of an *imagined* following of a rule. Thus the familiar pattern of moral argument, 'how would it be if everyone did that?' cannot have any effect on a consistent utilitarian unless his action really will have the effect of making everyone do it, which is usually pretty implausible. A purely imaginary consequence can no more figure in a utilitarian calculation, than the happiness or unhappiness of purely imaginary persons can. So the gas bill model, at least, cannot in itself get us on the road to that kind of generalization argument.

If the utilitarian wants to justify the use of the generalization argument which deals in imagined consequences, he will have to take a further step away from the actual consequences of particular choices, and deal in terms of the *actual* consequences of people's *thinking* in terms of *imagined* conse-

quences. But now he seems to be getting further and further away from the original utilitarian advantages. For, first, the supposed calculation of the utility of people's thinking in terms of imagined consequences—as against their working out particular cases; or working out the consequences of more specific rules; or taking the local morality as a going concern; or many other possibilities—this calculation begins to look more and more like bluff. How does he know what the consequences of these various practices might be? Second, he does know one thing, at least: that the more general the provision to which the utilitarian calculation is attached, the more cases there will be in which particular calculation in that case would have produced a different result, so the more tactical disutility he is licensing in his pursuit of strategic utility. In view of the doubtfulness of the strategic utility, this ought to worry him: one of the motivations of utilitarianism, after all, was a hardheaded injunction to think in terms of calculable consequences, and not just to rely on tradition, received practice and so forth.

The more one considers rule-utilitarianism, the more pressing this sort of point becomes. Turning once more to the gas bill model, we can recall that what principally made the uniform practice sensible was the cost of interfering with it. The analogy to this in ordinary moral deliberation is the disutility of calculating particular consequences. But the effect of that argument is cancelled out if we consider a case in which the particular calculation *has already been made*: and this is so in the morally disquieting

cases which presented the second kind of difficulty rule-utilitarianism was supposed to deal with. If calculation has already been made, and the consequences of breaking the rule are found better than those of keeping it; then certainly no considerations about the disutility of calculation could upset that result. And, indeed, it is very difficult to see how *anything*, for a consistent utilitarian, could upset that result. Whatever the general utility of having a certain rule, if one has actually reached the point of seeing that the utility of breaking it on a certain occasion is greater than that of following it, then surely it would be pure irrationality not to break it?

This consequence has indeed been drawn by some tough utilitarians, such as J. J. C. Smart. If utilitarianism can be got going at all, then I am sure theirs must be the right attitude to it: it is a special doctrine, not necessarily coincident with contemporary Western moral ideas in all respects, and one must expect it to have what may well seem unpalatable conclusions. Contrary to this, one feature of much modern utilitarian theory is that it is surprisingly conformist. Bentham and Mill regarded the Greatest Happiness Principle as an instrument of criticism and thought that by appeal to it they could show that many Victorian moral beliefs were mistaken and irrational, as indeed they were. But, except for the well-established areas of sexual and penal reform, themselves inherited from Bentham and Mill, modern utilitarian theorists tend to spend more effort in reconciling utilitarianism with existing moral beliefs than in rejecting those beliefs on

the strength of utilitarianism. One recent writer, for instance, has taken great and honest pains to show that public executions could not, as might seem, be justified on utilitarian grounds. He is left with some frank doubts; but these are doubts about the application and formulation of utilitarianism, and not, as they surely should be, doubts about whether public executions might not be reintroduced. This is an absurd case. But more generally all the many human qualities which are valued and yet resist utilitarian treatment, such as an unaccommodating passion for justice; certain sorts of courage; spontaneity; a disposition to resist such things as useful experiments on senile patients or the use of napalm on some people to secure (as it is supposed) the happiness of more people, often elicit from utilitarian theorists attempts to accommodate utilitarianism to those values rather than condemnation of such values as irrational legacies of a pre-utilitarian era. This is no doubt a tribute to the decency and imagination of those utilitarians but not to their consistency or their utilitarianism.

Rule-utilitarianism, as the enterprise of trying to hold on to something distinctively utilitarian, while knocking the rougher edges off it, seems to me a failure. This middle ground is not logically habitable. As opposed to this, one might, on the one hand, take the line already referred to, of Smart and others, and pursue act-utilitarianism modified only within the recognizable limits of the gas bill model. This is at least consistent with the misguided aims of utilitarianism, and the fact that it yields some distinctive

and (to many) unpalatable particular moral results should be a matter of no surprise. If, on the other hand, you desert this territory and start to apply the utilitarian principle to more and more general practices and habits of thought, what you come out with is unlikely to have any distinctively utilitarian content at all. This capacity of utilitarianism, once detached from the ground level, to annihilate itself, can be illustrated by a brief argument, with which I shall end. Its empirical premises are not, perhaps, beyond doubt. But they are certainly at least as plausible as most of those generally used by utilitarians in such matters.

One disturbing effect of people being active and conscious utilitarians is that it tends to debase the moral currency: a Gresham's Law operates, by which the bad acts of bad men elicit from better men acts which, in better circumstances, would also be bad. There is a simple reason for this: a utilitarian must always be justified in doing the least bad thing which is necessary to prevent the worst thing that would otherwise happen in the circumstances (including, of course, the worst thing that someone else may do)—and what he is thus justified in doing may often be something which, taken in itself, is fairly nasty. The pre-emptive act is built in to utilitarian conceptions, and certain notions of negative responsibility (that you are as responsible for what you fail to prevent, as much as for what you do) are by the same token characteristic of it. This being so, it is empirically probable that an escalation of pre-emptive activity may be expected; and the total con-

sequences of this, *by utilitarian standards themselves*, will be worse than if it had never started.

The utilitarian who is immersed in the system, however, cannot do anything about this; he must think in terms of actual consequences, and nothing in the realm of actual consequences (at least, nothing helpful) will now be effected by some gesture of principle—there is no way in which, from where he is, he can lead a dash to morally higher ground. Stepping back in reflection, however, he can consider how utilitarian aims might have been better realized than they have been in a world of utilitarians interspersed with villains. No doubt they would have been if there had been no villains: but that, certainly, is Utopian. What looks more hopeful is a state of affairs in which enough people are resistant to continuing the rot: resistant, for instance, by there being a range of things that they cannot consider doing, or bring themselves to do, or put up with being done, whatever other people do or may do. There is a limit to their pre-emptive activities. Enough people, enough of the time, it seems, have to be prepared to stick at doing various things, whatever the consequences may be. That means that enough people, enough of the time, do not have to think as utilitarians; they have, quite definitely, to think as non-utilitarians. Nor will it do for them to preserve at the back of their mind the utilitarian rationale in coexistence with the required moral bloody-mindedness. For they have to be able to resist utilitarian temptation in the most difficult circumstances, when much obvious harm will follow

from resisting it, and for that their non-utilitarianism has to be very deeply engrained.

Some utilitarians have reached, if not quite for these reasons, something rather like this conclusion, and have thought that what it showed was that the truth of utilitarianism could be known to a responsible élite, but should not be too widely spread among the masses. Such a proposal is both personally and socially hopeless. Personally, since the state of mind ascribed to the reflective utilitarian, and the attitude to others it involves could be honestly held, if at all, only by a very innocent man (as perhaps Sidgwick was), and no reflective man in our age can be that innocent. Socially, because the educational and other institutions required to embody such a view would have to be quite different from anything we could now expect or tolerate, or that utilitarianism itself could want.

If all this is true, then the world which the reflective utilitarian must finally settle on as most likely to yield the outcomes he wants, is a world in which the Gresham's Law is defeated because enough people enough of the time are deeply disposed against thinking in a utilitarian fashion. It is not possible that this disposition should coexist with believing in utilitarianism; nor is it acceptable or socially possible that most should have this disposition while others, the utilitarian élite, should believe in utilitarianism. All that is left is that the world which would satisfy the utilitarian's aspirations would be a world from which belief in utilitarianism as an overall

moral doctrine was totally absent, except perhaps as a minor and ineffective eccentricity.

So, if utilitarianism is true, and some fairly plausible empirical propositions are also true, then it is better that people should not believe in utilitarianism. If, on the other hand, it is false, then it is certainly better that people should not believe in it. So, either way, it is better that people should not believe in it.