

3 Utilitarianism and moral self-indulgence

My problem arises from the question of what one is to do in circumstances where there are strong reasons, particularly of a utilitarian kind, for doing something which one finds morally distasteful, and against which one has a strong personal commitment. It also of course involves questions of what one says and thinks about other people's actions in such situations. My concern is with a charge that can be brought against people who reject morally distasteful acts in such cases, namely that they are guilty of a certain kind of self-indulgence. When the agent's refusal takes the particular form of saying that while others, no doubt, will bring evil about, at least it will not come about through *him*, the charge may handily take the form of saying that the agent displays a possessive attitude towards his own virtue.¹

The problem particularly comes up in relation to utilitarianism. If the reasons for the act are, from a utilitarian point of view, strong enough, then utilitarians will say that the fact that the act is morally distasteful is certainly not an adequate reason against doing it in this case; as a general characteristic of acts of this sort, it is largely irrelevant to questions of what to do here and now, though it may be relevant to other aspects of the situation – thus we may think well of the agent for finding this kind of act distasteful, his reaction being taken as a reassuring sign of good character. It is in the context of a critique of utilitarianism that I have elsewhere² invoked the notion of *integrity* in this connexion, and it is in this context that I shall discuss the problem here, taking, that is to say, the reasons inviting one to the distasteful act as utilitarian reasons. However, the general structure of this problem for individual action is not confined to this sort of case, and I hope that my discussion will help to bring that out.

¹ The phrase appears in a discussion of these issues by Jonathan Glover, *Proc. Arist. Soc.*, Supplementary Volume XLIX (1975).

² In J. J. C. Smart and B. Williams, *Utilitarianism: For and Against* (CUP, 1973).

There is a set of problems very closely related to this one, which are problems of politics – taking this in a broad sense of action in a public capacity in a public domain, though the clearest and most important issues arise from matters of state. The clearest of all cases are actions by politicians in the exercise of their office in the context of international affairs, but similar issues can range down to such matters as rising politicians making deals to advance their careers, and their aspirations, and their supporters' moral hopes. These problems of political morality – the matter of 'dirty hands'³ – I shall not try to discuss here. It is important that they are not just a special case of the issue I shall discuss, not just examples of that issue arising in the political domain. In the clearest examples of the political, we have two special features. First, the agent stands in a relation to others – citizens, supporters, electors, etc. – in which he is supposed to effect results which involve, and can be known to involve, such acts; and this relation itself can have a moral dimension, for instance of trust. I say 'is supposed to effect results which involve...' rather than 'is supposed to perform such acts' because the public sometimes do take, and the media often pretend to take, a moralised view by which politicians are supposed not to do the acts required for what they are supposed to achieve. Second, the sphere of operation is itself less moralised and less structured by moralised expectations than at least a lot of other activities in at least settled communities: international relations are of course the prime example of this.

These two factors are different from one another. Issues of the second kind might arise even if there were no-one you were responsible to and for: some, though not all, traditional moralists have thought that there were restrictions on the extent to which moral considerations apply in the state of nature, and believed in the moral analogue to *silent leges inter arma*. The first feature, again, can arise without the second, but without the second factor, there would be greater doubt that the role being exercised by those responsible was a legitimate or acceptable role – the expectations people have of the leader are affected by their perception of the terrain over which he is leading them.

For these reasons, questions of dirty hands are not just a special case of the present problem: or rather, to assume that they are is to beg a major question about the answer to them. The present problem is about the nature and proper content of what is undoubtedly a person's

³ See Michael Walzer's discussion in *Philosophy and Public Affairs* (Winter, 1973); and chapter 4, below.

individual moral judgement, and (leaving aside an outlook which actually *defined* moral considerations in terms of utilitarianism) concerns what is certainly a moral choice between moral solutions. The question of dirty hands, at least in its strongest form, concerns what role a person's individual moral judgement is supposed to play in the business at all. The present problem is interested in the individual's moral consciousness and how it should appreciate the situation; the question of dirty hands raises the issue of whether his moral consciousness, and how it appreciates the situation, is not just an irrelevance.

One issue that does notably arise with both these questions, but which, again, I shall not discuss, is the extent to which, and the ways in which, actions offensive to morality can be retrospectively justified – perhaps even morally justified – by success; and what, if they can, may count as success. In its least palatable form, this is the view that even political atrocities can be justified by history. However, neither the unpalatableness of that application, nor (still less) some supposed guarantee offered by the sense of the term 'moral', should lead us to underestimate this view in general: it has more to it than people like to admit. But it is a topic for another occasion.

Our problem arises with cases in which the agent is faced with a reliable choice between a detestable action and an outcome which will be utilitarianly worse: where 'a reliable choice' means that he has a choice between doing and not doing the action, and it is certain beyond reasonable doubt that if he does not do the action, then that outcome, or something yet worse than that outcome, will follow.⁴ There are familiar arguments to suggest that no, or few, such choices are in fact reliable. On the one hand, utilitarians urge the importance of side-effects in calculating the balance of utility between acting and refraining: when side-effects are included, the detestable action will be said to possess less utility than at first appearance, and may have less utility than the alternative outcome. It is worth remarking that the level of probability attaching to these considerations is usually left quite indeterminate. Some of these effects, on which great weight in the abstract is put by defenders of utilitarianism, are so problematical that in any actual case a consistent and clear-headed utilitarian would be bound very largely to discount them. In any case, we shall assume that

⁴ It can be accepted, presumably, that the more horrible the action which is to be justified by the prospect of a given good outcome, the more probable it has to be that the outcome will indeed follow the action: suppose this already allowed for in the case. For two examples of the kind in question, see Smart and Williams, *op. cit.*, pp. 97–9.

we are dealing with a class of cases in which, when all these considerations are counted in, the balance of utilitarian advantage favours the (otherwise) detestable action. Clearly no utilitarian could say, and few would want to, that there could not be any such case.

An alternative tack for casting doubt on the reliability of such choices, used this time by anti-utilitarians, is to suggest that the efficacy of the detestable action (e.g. in preventing great harms which would otherwise occur) is more doubtful than the example supposes. This is a line often taken by those defending an absolutist position in cases of detestable actions extorted by threats made by hijackers and so forth, to the effect that the very character of the threat shows that one has reason to doubt the efficacy of giving in to it: why should one expect such threateners to keep their promises anyway? As a *general* line of argument, this seems to me, bluntly, a cop-out. Of course there are some cases in which it is a reasonable bet that nothing is to be gained by giving in to threats, but there are others in which it is not a reasonable bet, and it is merely an evasion to pretend that we have an *a priori* assurance, applicable to every case, that it is inadequately certain that the action will have its expected effect.⁵

In any case, there are only certain sorts of examples to which this line of reasoning can be relevantly applied at all, namely those in which, if the threateners fail to deliver, the all-round outcome is worse than if one had not done the detestable action. Not all cases which raise our problem – not even all that involve threateners – are of this structure. There is the case in which I am invited to kill one man, and told that if I refuse, someone else will shoot that man and several other men as well. If we think solely in terms of outcomes, then the only conceivable outcome actually better than those which involve my accepting, is that in which I refuse and they decide not to kill anyone; but there is absolutely no probability of that at all. If the other persons do what is analogous to promise-breaking in a hijacking case, namely that I accept and they nevertheless kill the rest, then the outcome, regarded as an outcome, is only the same as what it certainly will be if I refuse. So in terms of outcomes, we need only some non-zero

⁵ The underlying idea seems to be the *unity of the vices*, a psychologically unsound principle. A bizarre application of much the same idea is an argument adopted by P. T. Geach from M^cTaggart, to the effect that we could have no reason to believe in an unjust hell: the only ground for belief in hell being revelation, we should have no reason to regard as trustworthy the communications of a God wicked enough to run an unjust hell. See *Providence and Evil* (Cambridge 1977), pp. 134–6.

probability that they will do what they say for my acceptance to be rational.

In general, arguments of this kind seem only too ready to confuse the idea that some factor ignored in the example is possible, with the idea that it has some indeterminately high probability. In this, they notably resemble some arguments brought forward by their utilitarian opponents. No doubt the reason is the same: each in its own way is trying to find a consequentialist argument for some sentiment which does not have its roots in consequentialist considerations at all.

Let us then grant a reliable choice of the kind in question. Someone who knowingly takes the anti-utilitarian course in such a case might be open to the charge of being concerned with his own integrity or purity or virtue at others' expense. To use one phrase as a general label, though it might not always be the best phrase, let us call this the charge of moral self-indulgence. The first things I want to discuss are certain necessary conditions of such a charge being appropriate. In doing this, I shall assume that this charge is not, and is not intended to be, just trivially *equivalent* to a disapproving claim that someone, for reasons of the moral kind, knowingly acted in an anti-utilitarian way. I take it that an equivalence is not intended, since one who makes this charge in this connexion surely intends to commend the utilitarian solution to such cases, and hence indirectly the utilitarian system, by bringing non-utilitarian outlooks in certain of their applications under a charge which has some independent force, and which might already be recognised as an objection. It is this independent force which I shall try to uncover; and I shall, more particularly, assume that the charge of moral self-indulgence imputes a specific kind of *motive*.

It is, in fact, neither a sufficient nor a necessary condition of this charge's being appropriate that the agent knowingly does from a motive of the moral kind something which is counter-utilitarian. It is not sufficient, for consider the case of a man who, courageously doing what he takes to be his duty (or even just courageously), gets himself killed in the course of a counter-utilitarian project. He may be rash or foolish, but not, on the strength of this, morally self-indulgent: what contributes to this may possibly include the fact that he pays a high price himself. It is also not a necessary condition. It is possible for someone to be open to the charge of moral self-indulgence when the moral considerations which influence him are themselves utilitarian ones. Someone might incur this charge in certain cases (not all) who, for reasons of the general utilitarian welfare, left high and dry someone

who depended on him. If the man who refused to shoot when invited to by the threateners was keeping his hands clean from what the utilitarian would regard as ultimately unreal dirt, *this* man is keeping his hands clean from what, for the utilitarian, is real dirt.

What would encourage one to bring this charge against this man? One feature might be that he did not really seem to care about any particular other beneficiaries very much. This cannot mean just that there were no particularly identified beneficiaries about whom he cared – *that* would be the case with, for instance, a man who honourably acted in the interests of the unidentified inhabitants of an identified town, or, to take a more radical case, acted to prevent radiation hazard to future people. Nor will refinements on this thought get us to the nub of the charge; but the thought is suggestive of something which is much nearer the nub of the charge. One thing the thought can express is the suspicion that what the agent cares about is not so much other people, as himself caring about other people. He has an image of himself as a virtuous utilitarian, and this image is more important in his motivation than any concern for other persons, in particular that person for whom he is specially invited to show concern.

It is this type of *reflexive* concern which, I suggest, is significantly related to the charge of moral self-indulgence. It can arise with any moral motivation whatsoever. Thus a person may act from generosity or loyalty, and act in a counter-utilitarian way, and not attract the charge of moral self-indulgence, but that charge will be attracted if the suspicion is that his act is motivated by a concern for his own generosity or loyalty, the enhancement or preservation of his own self-image as a generous or loyal person. In the case of a man who acts in a counter-utilitarian way for reasons not of the moral kind, the charge of moral self-indulgence will not in any case stick, since 'moral' is not the sort of self-indulgence, if any, that he is going in for. But there are highly analogous contrasts in the matter of reflexivity. It is one thing for a man to act in a counter-utilitarian way out of his great love for Isolde, another for him to do so out of a concern for his image of himself as a great Tristan. The distinction applies even to the case of selfishness. One can act selfishly, that is to say, in a manner motivated by desire of things for oneself and indifferently to the welfare of others, but it would be different from that to act from a conception of oneself as a person who so acts. While the latter is unlikely to be nicer, it has a chance of being a bit grander.

I take it that there is in general a clear conceptual distinction between

the first-order motivation in each of these cases, and its reflexive second-order substitute. After that very general recognition, however, there are many respects in which even at the analytical level, let alone in psychological reality, boundaries are quite unclear. I shall make one or two remarks on what is obviously a large set of questions.

One necessary condition of ascribing the second-order motivation to an agent is that we also take him to possess the concept of the first-order motivation in question. A particularly clear distinction between the two types of motivation is available where it is possible to be motivated in a certain moral way without possessing the relevant concept of that motivation at all. Some types of virtuous motivation permit this, and it is one more mistaken consequence of Kantian moral theory that the only genuine moral motivation is taken to be one which essentially involves the agent's being conscious of that type of motivation. But even if an agent does possess the concept of a certain virtuous motivation, it may be that he does not apply it to his own case: in the space provided, with some virtues, by this possibility, there is room for such a thing as intelligent innocence. And even if, last, the concept is applied and the thought of his own disposition is present, that is not the same as his motivation being provided by that thought. It is a point worth further inquiry that in the case of some virtues (such as, perhaps, courage) the presence of such a thought may be encouraging to the first-order motivation, whereas with others it is not so, the presence of the thought tending to destroy the first-order motivation. To the extent that this latter is so, there will be a reason (there are others) why some virtues are only imperfectly accessible to highly self-conscious and reflective agents, as there are other virtues fully accessible only to them.

It may well be that the route to acquiring and sustaining the first-order virtuous motivations requires a kind of self-esteem which may involve to some degree and in some form second-order motivations. It is a question of psychological theory to what extent that is so, though that extent is certainly limited, for instance by the matters of concept-possession which have already been mentioned. It is a psychological matter also, less perhaps of theory than of common observation, to what extent what sort of motive actually operates. Nothing I say is meant to imply that it is in the least easy to tell to what degree what sort of motive is operating, in someone else's case or – what in the nature of this matter is a very different thing – one's own.

However, even if there is a difference between these sorts of

motivations, there remains a question about what, if anything, is supposed to be wrong with the second-order motivation – in particular, what about it makes it self-indulgent. Indeed, some philosopher might argue that for at least some kinds of second-order motivation there could not be anything wrong with it. For on the account given so far, it looks as though a man would be motivated in some such second-order way if he were to ask himself ‘What would I do if I acted as a generous man would act here?’, and were motivated to act on the answer; and if he gave the right answer to the question, and acted on it, then it looks as though he would do just what a generous man would do, and for no worse reason. Is that moral self-indulgence?

No; though as a picture of moral deliberation the pattern is surely very distorted (whether the distortion is in the picture or in what is pictured). What is lacking from this for it to be, however odd in other ways, a matter of self-indulgence is some element of self-esteem – a point suggested by the fact that it is, after all, the generosity of some hypothetical ideally generous person that is invoked here, not the agent’s own. Here we can be misled by phrases such as ‘he is concerned with being generous’. This may mean merely that he is concerned to do the generous thing in a sense in which that is what any generous man is concerned to do; or that he is concerned to conform his conduct to some paradigm of a generous man, like the agent just mentioned (this kind of reflexivity looks, in fact, like a familiar example of a more primitive, rather than a fuller, moral development); or it may mean that he is concerned with his own generosity, where this implies that he had substituted for a thought about what is needed, a thought which focuses disproportionately upon the expression of his own disposition, and that he derives pleasure from the thought that his disposition will have been expressed – rather than deriving pleasure, as the agent who is not self-indulgent may, from the thought of how things will be if he acts in a certain way, that way being (though he need not think this) the expression of his disposition.

It is this sort of reflexivity which invites the name ‘self-indulgence’. It involves a reversal at a line which I take to be fundamental to any morality or indeed sane life at all, between self-concern and other-concern; it involves a misdirection not just of attention, though that is true too, but genuinely of concern, and they both issue in differences in what actually gets done. Distortions which are due primarily to diverted attention, are familiar also with skills; those which come from diverted concern, the virtues share with the affections. These differences

in what gets done fit in with something noted earlier in the matter of courage, the evidential weight attached in these questions to the agent's himself paying a price; he *can* do that in the course of reflexively regarding his own virtue, but the space for it becomes more constricted.

These remarks about reflexivity and moral motivation involve a claim about a question which is hard and important and has been inadequately studied: how we are to picture the expression of moral dispositions in an agent's deliberative thought. We have some views in philosophy about the reference to dispositions in explaining and evaluating other people's conduct. We have some views about the occurrence of moral considerations in practical deliberation (though they are largely restricted to questions about the function in deliberation of 'moral judgements'). What we seem to lack is any coherent representation of something which is certainly true, that distinctive moral dispositions, such as generosity, are expressed in the content (and not just the occasions) of the agent's deliberations. The one claim that I make about that subject here is that the characteristic and basic expression of a moral disposition in deliberation is *not* a premiss which refers to that disposition – it is not the basic characteristic of a generous man's deliberations that they use the premiss 'I am a generous man'. Whatever one goes on to say about this subject, that negative claim is surely correct. Though the generous man is partly characterised by what goes into his deliberations, it is not that what goes into them are reflections on his generosity.⁶

We are now in a position to see better the relations between utilitarianism and integrity in the matter of moral self-indulgence. If the objectionable feature of moral self-indulgence is identified as a certain kind of reflexive motivation, then it cannot stand in any simple contrast with utilitarian motivation. For, first, it can be contrasted with many things other than utilitarian motivation – as, in general, with first-order virtuous motivations. Further, utilitarian benevolence is

⁶ Nor, we can add, is it merely thoughts such as 'he needs help'; the occurrence of such thoughts certainly mark out some men from others, but does little to mark out generous men from non-generous. Nor is it the 'moral judgement', 'I ought to help'; apart from well-known questions about the connexion of that with motivation, it is not specially the mark of a generous man to have or act on that thought. An answer will probably have to start from the idea that the basic representation in deliberation of such a disposition is in the form 'I want to help...'; this has the further advantage of not making it unintelligible how such moral considerations can be weighed in deliberation against quite different considerations.

itself open to this reflexive deformation. The reason why utilitarian motivation seems to many the unique enemy of moral self-indulgence is that it seems the purest expression of other-concern as opposed to self-concern – isn't utilitarianism just the expression of concern for everyone, among whom self is outnumbered by others? But in fact the distinction between other-concern and self-concern is in no way the same thing as the distinction between utilitarian and non-utilitarian, and in the sense in which other-concerned motivations which are not those of utilitarianism are capable of reflexive deformation into one kind of self-concern, so is utilitarian motivation itself.

What about concern with one's own *integrity*? The simplest thing to say about this would be that integrity is one case of a virtue, and that, like other virtues, it is subject to reflexive deformation. But I think that this would be wrong; rather, one should perhaps say that integrity is not a *virtue* at all. In saying that, I do not mean that there is not all that much to be said for it, as one might say that humility was not a virtue. I mean that while it is an admirable human property, it is not related to motivation as the virtues are. It is not a disposition which itself yields motivations, as generosity and benevolence do; nor is it a virtue of that type, sometimes called 'executive' virtues, which do not themselves yield a characteristic motive, but are necessary for that relation to oneself and the world which enables one to act from desirable motives in desirable ways – the type that includes courage and self-control. It is rather that one who displays integrity acts from those dispositions and motives which are most deeply his, and has also the virtues that enable him to do that. Integrity does not enable him to do it, nor is it what he acts from when he does so.

If that is right, we can see why integrity, regarded as a virtue, can seem to smack of moral self-indulgence. For if it is regarded as a motive, it is hard to reconstruct its representation in thought except in the objectionable reflexive way: the thought would have to be about oneself and one's own character, and of the suspect kind. If integrity had to be provided with a characteristic thought, there would be nothing for the thought to be about except oneself – but there is no such characteristic thought, only the thoughts associated with the projects, in carrying out which a man may display his integrity. Relatedly, one cannot directly bring someone up to possess integrity, in the sense of teaching him to display or exercise it; rather one brings it about that he genuinely cares for something and has the characteristics necessary to live in the spirit of that.

But what of the thought ‘not through me’ – the thought that even if others are going to bring evil and injustice into the world, it will not be by my agency that it comes about?⁷ This, certainly, is already a reflexive thought, and involves at least one step away from the simply unselfconscious expression of counter-utilitarian dispositions. The thought, however, is not in itself a motivating thought, and those words do not express any distinctive motivation. It is not merely that they do not on all occasions express some one motivation. Rather, they do not, in themselves, express any motivation at all: if one is motivated *not to do it oneself*, then there is some (other) motive one has for not doing it. One such motive is fear, and in the particular form, perhaps, of the fear of pollution, it can attract the accusation of cowardice to some agents in the sort of circumstances we are discussing. With the motivation of fear in general, it is often the case that the agent *would prefer to be able to do* whatever it is he fears doing. In relation to that, the fear of pollution is a special case, providing either an exception to it, or a peculiarly complex instance of it. But in any case, fear, of whatever kind, is by no means always the motive of agents who use those words.

A quite different, perhaps limiting, case of a motive lying behind those words would be one related to pride, the motive of one of whom it is not true that he would prefer to be able to do it – he could do it – but who does not want it done, and refuses to be made to do it by another’s providing him with reasons for doing it. A bare, unsupported motive of this kind could hardly be adequate to the cases we have in mind – because the interests of innocent parties have been thrown into the reasons for acting, this would be, too much, arbitrary self-assertion. But a similar, though different, thought can be expressed by the agent in our case: similar, in that he registers a refusal to be coerced by the threats, inducements or example of others; different in that he is not just asserting his own independence and right to refuse, but expressing the other motivations he has for not doing the act in question.

Utilitarians will, or course, dispute his right to refuse, but the point is that the agent’s affirmation ‘not through me’ does not, in such a case, express a motivation of the suspect, ‘self-indulgent’, kind. In itself, it does not represent any motivation at all, and the motivations which

⁷ Glover, *op. cit.*, has called this the ‘Solzhenitsyn principle’, after a passage in that writer’s Nobel oration. The name is well invoked; but this thought should not be regarded as a *principle*.

can lie behind it include some which are, for various reasons, suspect and some which are not. The reflexivity of the utterance does not represent in itself any suspect motive, but only the self-consciousness of the refusal, however the refusal is motivated.

There are many and various forms of dispositions, patterns of feeling and desire, which can motivate people to counter-utilitarian acts; some themselves virtues, some more particular projects, affections and commitments. The question I turn to last is the place that utilitarianism can allow to such dispositions. They can be variously admired or deplored, cultivated or discouraged. Some may indeed be admired and encouraged for what are, remotely and ultimately, utilitarian reasons, in the sense that human welfare is served indirectly by the presence of these dispositions in the world. I think that it is wrong to try to reduce all questions of the assessment of such dispositions to utilitarian considerations, and indeed that it is incoherent, since there is no coherent view of human welfare itself which is independent of such issues as what people care for, in non-utilitarian spirit, with regard to such things as these dispositions. But that is not my concern here, and if the present argument goes through for those dispositions of this type which can be granted indirect utilitarian value, then it will presumably have some *ad hominem* force against utilitarianism.

The difficulty is that such dispositions are patterns of motivation, feeling and action, and one cannot have both the world containing these dispositions, and its actions regularly fulfilling the requirements of utilitarianism. If you are a person who whole-heartedly and genuinely possesses some of these admirable dispositions, you cannot also be someone in whose thought and action the requirements of utilitarianism are unfailingly mirrored, nor could you wish to be such a person. If you want the world to contain generous, affectionate, forceful, resolute, creative and actually happy people, you do not wish it to contain people who uniformly think in such a way that their actions will satisfy the requirements of utilitarianism.

The supposition that one might combine the dispositions one wants and admires in the world with actions that maximally satisfy the utilitarian criterion stems from a number of errors. One is an idea, which utilitarianism, though it denies it, is in fact disposed to share with other pictures of moral experience, and indeed of practical rationality in general, that the processes of practical thought are transcendental to experience and do not actually take up any psychological room. But in fact to think in one way rather than another about

what to do is to be empirically different, to be a certain kind of person, and it is not possible to combine all kinds of reflection with all kinds of disposition. Utilitarians neglect this to some extent at the level of the individual, but they have made a speciality out of neglecting it at the social level, supposing for instance that there could be an élite of utilitarian thinkers who possessed an esoteric doctrine unknown to others, without there being specified any form of social organisation to make this structure a social reality.

Second, there is the error, also shared with others, of dissociating moral thought and decision from moral feeling. It is a commonplace that there is a form of weakness which consists in being overcome by unstructured moral feeling and there is another which consists in a kind of squeamishness. These are often failures of confusion, of lack of self-knowledge. But the cure for them cannot or should not consist in teaching people to discount their moral feelings, to dissociate themselves from them.⁸ Theorists who encourage this are fond of such cases as that of the lapsed and now unbelieving Catholic who feels guilty when he does not go to Mass. But whatever is to be said about that case, it cannot be a paradigm of what the utilitarian needs. The lapsed Catholic aims to dissociate himself entirely from the Mass and its claims, to reach a position from which no such feeling has any significance at all. But no such thing is true of the man involved in counter-utilitarian feelings in a case such as we are discussing. These feelings represent something he in general stands by, and which the utilitarian, we are supposing, wants him in general to stand by.

No one is suggesting that moral feeling should express itself unmodified by thought (at the limit, this is not even a comprehensible idea). There are, further, some moral feelings, particularly concerned with the observance of rules, which can be formed by experience in ways which to some extent fit round and accommodate utilitarian thoughts: it is so, up to a point, with the rules of promise-keeping and truth-telling. But there is no reason at all to believe, for many dispositions of the kind that it is desirable to have in the world, that

⁸ A theory of the moral sentiments is needed here. One approach to the questions of dissociation from moral feeling might be suggested by a certain contrast between moral feeling and sense-perception. Those views, of rationalist type, which most strongly advocate dissociation from perceptual sensations, at least emphasise a truth, that the aim of objective knowledge is to dissociate thought about the world, certainly from what is distinctively oneself, and perhaps (on realist views) from anything that is distinctively human. But that cannot be the aim of moral thought and experience, which must primarily involve grasping the world in such a way that one can, as a particular human being, live in it.

they can retain their position and significance and yet systematically make way, whenever required to, for the deliverances of utilitarian thought, the feelings associated with the disposition being made the objects of dissociation.

Relatedly, there is not much to be got out of a third line of thought, which can also encourage an oversimple view of these problems, the supposedly clear distinction between judging the act and judging the agent. If a man has a disposition of a kind which it is good that he has, and if what he did was just what a man with such a disposition would be bound to do in such a case, but (as I claim must sometimes be so) was counter-utilitarian: what is the force of saying that what he did was as a matter of fact wrong? It is important that it does *not* have the force – which really would give some point to the distinction of act and agent – that, if he had been in a position to conduct his deliberations better, he would have acted differently. He conducted his deliberations as such a man does, and it is good that he is such a man. By the same token, it cannot mean that we ought to try to bring people up to be such that they do not make such mistakes. If there is any content to saying that this man did the wrong thing, it must be compatible with our thinking that it is a good thing that people do not always do the right thing; and not just in the very general sense in which we may reflect on the uncovenanted benefits which can flow even from dire acts, but in the more intimate sense that we want the world to contain people who when they ask themselves ‘what is the right thing to do?’ will, on definitely specifiable sorts of occasion, give the wrong answer.

The utilitarian’s theory, once he admits the value of these dispositions, takes the question ‘what is the right thing to do?’ a long way away from the question ‘what answers is it desirable that people should be disposed to give to the question “what is the right thing to do?”?’ The tension created by this separation is very great, and there is very strong pressure, if utilitarianism is to retain any distinct identity *within* moral thought, for it to reject or hopelessly dilute the value of these other dispositions, regressing to that picture of man which early utilitarianism frankly offered, in which he has, ideally, only private or otherwise sacrificeable projects, together with the one moral disposition of utilitarian benevolence. I hope to have shown that that false picture cannot be commended to us by rejecting other moral motivations, in their counter-utilitarian appearances, as pieces of moral self-indulgence.