I Utilitarian Accounts: State of Mind or State of the World?

How are we to understand 'well-being'? As 'utility', say the utilitarians, aware that this technical term itself needs explaining.[1] What is 'utility'? 'Pleasure and the absence of pain', the classical utilitarians said,[2] not realizing how much the words 'pleasure' and 'pain', especially in the stretched sense they attached to them, needed explaining.

Two main traditions about 'utility' have grown up. One sees it as a state of mind, the other as a state of the world. Is 'utility' mental states (e.g. pleasure, pain) or states of the world which fulfil desires (e.g. economists' 'preference')? If mental states, is it only one sort, or many? If many, what links them? If fulfilment of desires, desires as they happen to be, or in some way improved? If improved, how? We can forget morality for the moment. Utilitarians use our rough, everyday notion of 'well-being', our notion of what it is for a single life to go well, in which morality may have a place but not the dominant one. This does not mean that our job is merely to describe the everyday use. It is too shadowy and incomplete for that; we still have to be ready for stipulation.

1. Mental State Accounts

When some utilitarians have spoken of mental states such as pleasure and pain, they have meant these terms so widely that their accounts get very near desire accounts of 'utility'. So we cannot always take this verbal difference as marking any real difference.

Still, the difference is often real enough. Bentham and Mill are, with ample reason, taken to be offering a psychological

account of 'utility'. Pleasure or happiness is presented as a 'state of feeling', and pain or unhappiness as a feeling on the same scale as, and the opposite of, pleasure or happiness. And the utilities of all our experiences are supposed to be determinable by measuring the amount of this homogeneous mental state that they contain.

The trouble with thinking of utility as *one* kind of mental state is that we cannot find any one state in all that we regard as having utility—eating, reading, working, creating, helping. What one mental state runs through them all in virtue of which we rank them as we do? Think of the following case. At the very end of his life, Freud, ill and in pain, refused drugs except aspirin. 'I prefer', he said, 'to think in torment than not to be able to think clearly'.[3] But can we find a single feeling or mental state present in both of Freud's options in virtue of which he ranked them as he did? The truth seems, rather, that often we just rank options, *period*. Some preferences—Freud's seems to be one—are basic. That is, preferences do not always rest upon other judgments about the quantity of some homogeneous mental state found in, or produced by, each option. When, in these cases, one speaks of one thing's yielding greater satisfaction than another, this seems best understood as saying that having the first is the fulfilment of a greater desire than having the second would be. One wants the first more than the second. But these desires are not ranked by independent quantities of satisfaction.

So, if the mental state account takes this simple form, the objections to it are insurmountable. And if we do not want to go over to a desire account, there are two ways we might now move. We might accept that utility is not one mental state but many, and then look for an explanation of how they are linked. Or we might, on the other hand, decide that utility is neither a matter of mental states nor of desire-fulfilment but of something, in a way, in between; we might say that it is a matter of finding enjoyment in various things, where 'enjoyment' is what we might call an attitude, or a conscious state, or a state of a person. I want to leave the second move until later; it is not easy to grasp, and it will be easier after we have looked at both the mental state and the desire accounts. So let us go back to the first move.

2. Sidgwick's Compromise

Suppose we said that utility consisted of several different mental states. What then would make them into a set? The obvious candidate would be desire; we could say, following Henry Sidgwick in borrowing something from each of the competing accounts, that utility combines a psychological element and a preference element. 'Utility', we could say, is 'desirable consciousness', meaning by 'desirable' either consciousness that we actually desire or consciousness that we would desire if we knew what it would be like to have it.[4]

The trouble with this eclectic account is that we do seem to desire things other than states of mind, even independently of the states of mind they produce. This is the point that Robert Nozick has forcefully made with some science fiction.[5] Imagine an experience machine programmed to give you any experience you want; it will stimulate your brain so that you think you are living the most ideal life, while all the while you float in a tank with electrodes in your brain. Would you plug in? 'What else can matter to us', Nozick asks, 'other than how our lives feel from the inside?' And he replies, surely rightly, that we also want to *do* certain things, to *be* certain things, and to be receptive to what there is in life beyond what humans make. The point does not need science fiction; there are plenty of examples from ordinary life. I certainly want control over my own fate. Even if you convince me that, as my personal despot, you would produce more desirable consciousness for me than I do myself, I shall want to go on being my own master, at least so long as your record would not be much better than mine. I prefer, in important areas of my life, bitter truth to comfortable delusion. Even if I were surrounded by consummate actors able to give me sweet simulacra of love and affection, I should prefer the relatively bitter diet of their authentic reactions. And I should prefer it not because it would be morally better, or aesthetically better, or more noble, but because it would make for a better life for me to live. Perhaps some such preferences, looked at with a cold eye, will turn out to be of dubious rationality, but not all will. This fact presents a serious challenge to the eclectic account of utility. If not all desirable things are mental states, yet they

matter to our well-being, the eclectic account is fissile. Which part of it should we retain: desire or mental states? It is hard to retain mental states, for if we did, we should then have, puzzlingly, to accept that when, with eyes wide open, I prefer something not a mental state to a mental state and so seem to value the former more than the latter, I get greater utility from what I value less.[6] Of the two, it is better to retain *desire*.[7]

Of course, 'mental state' is a vague expression. Perhaps Sidgwick and others use it broadly enough to include, say, knowledge. However, that does not seem to be Sidgwick's intention, and in any case it would still not be broad enough. I also want to *be* my own master, and it would take more broadening to include that. It seems more promising to abandon 'mental state' altogether and to try defining 'utility' solely in terms of desire: utility consists, we might try saying, in the fulfilment of desire.

3. The Actual-Desire Account

The simplest form of desire account says that utility is the fulfilment of *actual* desires. It is an influential account. Economists have been drawn to it because actual desires are often revealed in choices and 'revealed preferences' are observable and hence a respectable subject for empirical science.[8] Also the same account of utility can then do service in both moral theory and theory of action; explanation of action has to appeal to what we in fact want rather than to such ideal notions as what we ought to want or would want if well-informed. And both philosophers and social scientists have been powerfully drawn to it because it leaves no room for paternalism; if actual desires determine distributions, consumers are sovereign and agents autonomous.

Yet, notoriously, we mistake our own interests. It is depressingly common that when even some of our strongest and most central desires are fulfilled, we are no better, even worse, off. Since the notion we are after is the ordinary notion of 'well-being', what must matter for utility will have to be, not persons' actual desires, but their desires in some way improved. The objection to the actual-desire account is overwhelming.

In any case, considerations of autonomy are, on reflection, no recommendation of it. Well-being and autonomy, no doubt, both matter morally. It is even likely that living autonomously would be part of any enlightened person's conception of a good life. But it just confuses two quite different ideas to adopt the actual-desire account of well-being just because it makes autonomy prominent. One consideration to keep in mind is that the question, 'What is the best account of "utility"?', should be kept distinct from the question, 'What is the account of "utility"—perhaps highly artificial and *ad hoc*—that yields a one principle, utility maximizing, moral theory that comes closest to adequacy?' It is wrong to try to build into the notion of 'utility' all the restrictions that morality needs, if they fit more naturally elsewhere in the theory.

4. The Informed-Desire Account

At this point, an obvious move is to say that desires count towards utility only if 'rational' or 'informed'. 'Utility', we might try saying, is the fulfilment of desires that persons would have if they appreciated the true nature of their objects. But we shall have to tone this definition down a bit. Although 'utility' cannot be equated with actual desires, it will not do, either, simply to equate it with informed desires. It is doubtless true that if I fully appreciated the nature of all possible objects of desire, I should change much of what I wanted. But if I do not go through that daunting improvement, yet the objects of my potentially perfected desires are given to me, I might well not be glad to have them; the education, after all, may be necessary for my getting anything out of them. That is true, for instance, of acquired tastes; you would do me no favour by giving me caviar now, unless it is part of some well-conceived training for my palate. Utility must, it seems, be tied at least to desires that are actual when satisfied. (Even then we should have to stretch meanings here a bit: I might get something I find that I like but did not want before because I did not know about it, nor in a sense want now simply because I already have it; or I might, through being upset or confused, go on resisting something that, in some deep sense, I really want.) It

is hard to get the balance between actual and informed desires quite right. But, to be at all plausible, the informed-desire account has to be taken to hold them in a balance something like the one I have just sketched.

The move to 'informed-desires' marks the first important break with the classical utilitarian tradition (we shall see several more in the course of the discussion). Bentham and Mill used 'utility' both to explain action and to set a moral standard; they used its empirical role in arguing for its moral role. But now 'utility' has taken on a shape to fit it for a normative role (it need not be only in moral theory; it could also be in an account of one person's well-being or an account of practical reason), and it is of doubtful relevance to a purely empirical account of motivation. So this account of 'utility' should no longer be seen to be attached, except historically, to certain theories of action. It is not committed to the view that action is the result purely of a vector of desire-forces. It is not committed to any Hume-like account of the role of reason and desire. We can no longer use historical connections as a guide to theoretical connections.

The informed-desire account starts with the recognition that actual desires can be faulty. What sorts of fault matter? Obviously, for one, lack of information. Some of our strongest desires rest on mistakes of fact. I make my fortune, say, only to discover I am no better off because I was after people's respect all along and mistakenly thought that making a fortune would command respect. Or I want an operation to restore me to health, not realizing that some pill will do just as well. What matters is the ultimate, not the immediate, object of my desire, and factual mistakes creep into matching the one to the other.[9] Or I develop one set of material desires not realizing that they are the sort that, once satisfied, are replaced by another set that are just as clamorous and I am no better off. The consumer-desires at the centre of the economists' stage can be like that.[10] Then another relevant fault is logical mistake. A lot of practical reasoning is about adapting means to ends and, like any reasoning, it can be confused, irrelevant, or question-begging. Then there are subtler faults.[11] Sometimes desires are defective because we have not got enough, or the right, concepts. Theories need building which will

supply new or better concepts, including value concepts. For instance, it is easy to concentrate on desires to possess this or that object, at the cost of the more elusive, difficult-to-formulate, desires to live a certain sort of life. And it is almost impossible to strike the right balance between the two main components of happiness—on the one hand, the discontent that leads to better and, on the other, contentment with one's lot.[12] One needs more than facts and logic to sort those problems out: one needs insight and subtle, perspicuous concepts. And with information, more is not always better. It might cripple me to know what someone thinks of me, and I might sensibly prefer to remain in ignorance.[13] What seems most important to the informed-desire account is that desires have a structure; they are not all on one level. We have local desires (say, for a drink) but also higher order desires (say, to distance oneself from consumers' material desires) and global desires (say, to live one's life autonomously). The structure of desires provides the criterion for 'informed' desire: *information* is what advances plans of life; information is *full* when more, even when there is more, will not advance them further. So there is only one way to avoid all the faults that matter to 'utility': namely, by understanding completely what makes life go well.[14]

This brings out another break with classical utilitarian tradition. Bentham, Mill, and Sidgwick all saw utility as having to enter our experience. But we desire things other than states of mind; I might sometimes prefer, say, bitter truth to comforting delusion. The informed-desire account has the advantage of being able to accommodate such desires. But the desire account does this by severing the link between 'fulfilment of desire' and the requirement that the person in some way experience its fulfilment, dropping what we might call the Experience Requirement.[15] If the delusion is complete, one believes that one has the truth; the mental states involved in believing something that really is true and believing a successful deception are the same. Or if a father wants his children to be happy, what he wants, what is valuable to him, is a state of the world, not a state of his mind; merely to delude him into thinking that his children flourish, therefore, does not give him what he values. That is the important point; the

informed-desire account does not require that fulfilment of desire translates itself in every case into the experience of the person who has the desire, and that is what gives the account its breadth and attraction as a theory of what makes life valuable. This seems to me the way that the informed-desire account has to develop. The definition itself is short: 'utility' is the fulfilment of informed desires, the stronger the desires, the greater the utility. The way that the account develops, however, shows that all of those key terms are to a fairly large degree technical.

(a) 'Desire'.In the present technical sense, desires clearly do not have to have felt intensities; they need not be linked exclusively with appetitive states (some are, but others are aims we adopt as a result of understanding and judgement); they need not have existed before fulfilment. Rather, desiring something is, in the right circumstances, going for it, or not avoiding or being indifferent to getting it.

(b) 'Informed'.In its technical sense, 'informed'[16]is the absence of all the faults that I listed just a moment ago. There is a historically important account of practical reason that goes roughly like this: reason alone can never determine action. The end of action must be something fixed on, in its own reasonless way, by desire; we reason, but deliberation is only of means.[17] It is hard to see what is at issue between those who say, with Hume, that reason alone cannot supply a motive and those who say, with Kant, that it can. But those of the latter pursuasion are right to this extent: in deciding how to act, we must try to understand what properties things and states of affairs have, and we must put our desires through a lot of criticism and refinement to reach this understanding.[18] In this sense, deliberation may be of ends, and important deliberation often is. So an 'informed' desire is one formed by appreciation of the nature of its object, and it includes anything necessary to achieve it.[19]

(c) 'Fulfilment'.Being 'fulfilled' cannot be understood in a psychological way, or we should be back with mental state accounts. A desire is 'fulfilled' in the sense in which a clause in a contract is fulfilled: namely, what was agreed (desired) comes about.

(d) 'Strength'.'Strength of desire' has several senses,

appropriate to different theoretical settings. The 'strongest' desire can be the winner, or it can be the most intensely felt. But strength of desire, in its technical sense here, has to be understood in connection with the structure that informed desires have. One does not most satisfy someone's desires simply by satisfying as many as possible, or as large a proportion. One must assess their strength, not in the sense of felt intensity, but in a sense supplied by the natural structure of desire. The desires I feel most intensely could be satisfied by your constantly imperilling my life and saving me only at the last moment,[20] whereas I should clearly prefer peace to peril; anyway, felt intensity is too often a mark of such relatively superficial matters as convention or training to be a reliable sign of anything as deep as well-being. That I prefer peace to peril suggests that global desires provide, in large part, the relevant notion of strength of desire: I desire the one form of life more than the other. True, sometimes we form global desires only on the basis of having summed local desires (for example, the global desire for a way of life based on a reckoning that day-to-day pleasures will be maximized that way). But even then we must rank that way of life against others that it excludes, and our preference between them will, it seems, be basic—that is, a global judgment not based on any other quantitative judgments. This means that the relevant notion of aggregation cannot be simply that of summing up small utilities from local satisfactions; the structure of desires already incorporates, constitutes, aggregation. It means also that the relevant sense of 'strength' is not simply the desire that wins out in motivation. If my doctor tells me that I shall die if I do not lay off drink, I shall want to lay off it. But I may later crack and go on a binge, and at that point my desire to drink will, in a perfectly clear sense, be strongest. If strength were interpreted as motivational force, then 'utility' would lose its links with well-being; what would be good for me would then be fulfilment not of my informed desires but of what I 'ought to desire' or 'have reason to desire'. So to retain the links with well-being, the relevant sense of 'strength' has to be, not motivational force, but rank in a cool preference ordering, an ordering that reflects appreciation of the nature of the objects of desire.[21]

5. Troubles With the Informed-Desire Account

There are strong objections to such an account. Is it even intelligible?[22] If our desires never changed with time, then each of us would have a single preference order, by reference to which what most fulfilled his desires over the course of his life could be calculated. However, life is not so simple; preferences change, and not always in a way that allows us totally to discount earlier ones. Suppose that for much of his life a person wanted his friends to keep him from vegetating when he retired but, now that he is retired, wants to be left to vegetate. Is there any intelligible programme for weighing desires that change with time and hence for maximizing fulfilment? If not, we may be driven back to a happiness or mental state account.

Yet all the problems that we have just seen with mental state accounts remain; defects in one account do not obligingly disappear with the appearance of defects in another. How do we determine how happy a person is? Is happiness a single mental state? If many, how are they linked? Mental state accounts are hardly a refuge from troubles. Moreover, there may be an acceptable programme for handling cases where preferences change with time. The notion of an informed desire needs still further development and may eventually be able to supply the weighting of desires that we need in these troublesome cases. Has our retired friend simply forgotten the satisfactions of a busy life? If so, his later desire has much less weight. Is it just a change in taste, on the model of no longer liking ice cream? If so, his earlier desire has much less weight. We shall have to come back to these problems when we discuss measurement, but for now I have to be content with suggesting that the prospects of making the informed-desire account work are certainly not less rosy than those of making a mental state account work.[23]

The other troubles are much more worrying. The breadth of the account, which is its attraction, is also its great flaw. The account drops the Experience Requirement, as we called it. It allows my utility to be determined not only by things that I am not aware of (that seems right: if you cheat me out of an inheritance that I never expected, I might not know but

still be worse off for it), but also by things that do not affect my life in any way at all. The trouble is that one's desires spread themselves so widely over the world that their objects extend far outside the bound of what, with any plausibility, one could take as touching one's own well-being. The restriction to *informed* desire is no help here. I might meet a stranger on a train and, listening to his ambitions, form a strong, informed desire that he succeed, but never hear of him again. And any moderately decent person wants people living in the twenty-second century to be happy and prosperous. And we know that Leonardo had an informed desire that humans fly, which the Wright brothers fulfilled centuries later.[24] Indeed, without the Experience Requirement, why would utility not include the desires of the dead? And would that not mean the account had gone badly awry? And if we exclude these desires that extend beyond the bounds of what affects well-being, would we not, in order to avoid arbitrariness, have to reintroduce the Experience Requirement, thereby losing the breadth that makes the informed-desire account attractive? The difficulty goes deep in the theory. In fact, it goes deep, one way or other, in any account of well-being.

Another attraction of the account is that desires have to be shaped by appreciation of the nature of their objects. Without that restriction, the account is not even a starter. But with it, do desires even matter any longer? It may be somewhat too simple to say that things are desired because valuable, not valuable because desired. Yet the informed-desire account concedes much of the case for saying so. What makes us desire the things we desire, when informed, is something about them—*their* features or properties. But why bother then with informed desire, when we can go directly to what it is about objects that shape informed desires in the first place? If what really matter are certain sorts of reason for action, to be found outside desires in qualities of their objects, why not explain well-being directly in terms of them? It does not seem that it is fulfilled desire that is the basis of well-being, but certain of its objects. And that points us, depending on what we decide those objects are, either back towards mental states or beyond utility altogether.[25]

6. Is There Something Between Mental State and Desire Accounts?

Mental state accounts are too narrow, desire accounts too broad. We ought to look in between.

I said a while ago that one way of correcting the flaws in the simple mental state account would be by making utility neither a mental state nor desire-fulfilment. I then postponed taking up that suggestion until after we had looked at desire accounts, and it is time to return to it now.

There is a cluster of terms which even in their everyday use seem to fall conveniently between mental states and fulfilment of desires: namely, enjoying or liking things, finding them pleasing or satisfying or fulfilling, being pleased or happy with them. Let us use the term *enjoyment* to cover them all. *Enjoyment*, in its ordinary use, is not anything so narrow as experiencing a single mental state or one of a range of states, but let us explicitly make it part of the present use of *enjoyment* that it is not. In similar spirit, let us specify that it is nothing so broad as merely having desires (even informed ones) fulfilled. Also, let us allow that people *enjoy* things other than states of mind; in fact, to treat the account sympathetically we should make the range of objects of *enjoyment* wide—wide enough to include, say, helping others or advancing knowledge. But let us put a limit to the range by requiring that all the objects of *enjoyment* fall within our experience. So we both push the boundaries out beyond its ordinary use and, at the same time, limit the expansion by reinstating the Experience Requirement.[26]

The enjoyment account will need a lot more work to make it clear. Is it just a desire account with the Experience Requirement tacked on? Does the notion of *enjoyment* do any work of its own? *Enjoying* and *liking*, in their ordinary uses, are closely connected with desire: the acid test for whether I enjoy or like something is whether, other things being equal, I go for it, or do not avoid it, or am not indifferent to getting it. *Enjoyment* in its technical sense is even more closely connected with desire. Many everyday uses of *like* and *enjoy* suggest a certain psychological tone, but that disappears in the technical sense. However, if *enjoying* is just having a favourable attitude, the revival of the Experience Requirement runs the risk of

being *ad hoc*. I have favourable attitudes towards many things that do not enter my experience. Why do they not count too?

Still, the enjoyment account is clear enough for it to face one serious trouble. It is attractive because, with the revival of the Experience Requirement, it restrains over-wide desire accounts. The intention is right, but the particular restraint applied seems not. It seems in the end simply too drastic. It bans things that our ordinary notion of well-being cannot, without damage, do without. It is common that, as many persons' values mature, such things as accomplishment and close authentic personal relationships come more and more to fill the centre of their lives. If the Experience Requirement excludes these values from 'utility', then 'utility' will have less and less to do with what these persons see as making their own lives good. And those values do seem excluded. Suppose that someone is duped into thinking that those close to him are behaving authentically. What enters his experience is the same whether he has the real thing or a successful deceit. But it is only the real thing, he thinks, that makes his life better. According to the enjoyment account, what affects well-being can only be what enters experience, and the trouble is that some of the things that persons value greatly do not. My truly having close and authentic personal relations is not the kind of thing that can enter my experience; all that can enter is what is common to both my truly having such relations and my merely believing that I do. And this seems to distort the nature of these values. If I want to accomplish something with my life, it is not that I want to have a *sense* of accomplishment. That is also desirable, but it is different from, and less important than, the first desire. And if I want to accomplish something, it is not necessary that I want my accomplishment to enter my experience—say that I know about it. That too is desirable, but it is still not the first desire. If either I could accomplish something with my life but not know it, or believe that I had but not really have, I should prefer the first. That would be, for me, the more valuable life. 'Valuable life', of course, is full of ambiguity. It can mean a life that is valuable because of its value to other persons. It can mean a morally valuable life, or an aesthetically valuable one, or one valuable in terms of some code, such as a code of chivalry. But my

ground for preferring the first sort of life would not be any of these; I should prefer it because it would be, considered on its own, considered simply as a life I must lead, a more fulfilling one. So it is a value that has to be found a place within the bounds of 'utility'.

The enjoyment account, too, has its serious troubles. To my mind, the best prospect for a utilitarian account of well-being is to hold on to the over-wide desire account and look for good reasons to rein it in.[27] It is harder to correct the over-narrow accounts. However, developing the desire account is a large job, which will be the subject of the next chapter.

II Utilitarian Accounts: The Desire Account Developed

1. How May We Restrict the Desire Account?

The informed-desire account will have to be abandoned unless we can find a way to restrict the desires that count. But we cannot do it with the Experience Requirement. That is where the last chapter left us.

The trouble, you will recall, comes from examples like these: I want the sympathetic stranger I meet on the train to succeed; I want people in the twenty-second century to prosper; Leonardo wanted humans to fly. All of them informed desires, but (the trouble is) their fulfilment not part of well-being.

The notion we are after is not the notion of value in general, but the narrower notion of a life's being valuable solely to the person who lives it. And this must itself impose restrictions on which desires count. As these examples show, the desires that count have to enter our lives in a way beyond just being our desires. So what we need to do is to make clear the sense in which only certain informed desires enter our lives in this further way. Think of the difference between my desire that the stranger succeed and my desire that my children prosper. I want both, but they enter my life in different ways. The first desire does not become one of my aims. The second desire, on the other hand, is one of my central ends, on the achievement of which the success of my life will turn. It is not that, deep down, what I really want is my own achievement, and that I want my children's prosperity only as a means to it. What I want is *their* prosperity, and it distorts the value I attach to it to make it only a means to such a purely personal end as my own achievement. It is just that their prosperity also becomes part of my life's being successful in a way that the prosperity of the stranger on the train does not.

But that can be only part of the story. It is not that informed desires count only if they become the sort of aims or goals or aspirations on which the success of a life turns. Good things can just happen; manna from heaven counts too. So we should try saying, to introduce more breadth, that what count are what we aim at and what we would not avoid or be indifferent to getting. What counts for me, therefore, is what enters my life with no doing from me, what I bring into my life, and what I do with my life. The range of that list is not so great as to include things that I cannot (e.g. the prosperity of our twenty-second-century successors) or do not (e.g. the sympathetic stranger's success) take into my life as an aim or goal. And Leonardo's wanting humans to fly would not count either; to the extent it became an aim of his life it was unsuccessful, and to the extent it was merely a wish it does not count.[1]

In a way the account is now circular. I appeal to our rough notion of well-being in deciding which informed desires to exclude from this account of well-being. But that, I think, does not matter. If what we were doing were taking a totally empty term, 'well-being', and stipulating a sense for it, then we could not, in the middle of the job, appeal to 'well-being'. But our job is not that. The notion of 'well-being' we want to account for is not empty to start with; utilitarians use our everyday notion, and our job is to make it clearer. So we are free to move back and forth between our judgments about which cases fall inside the boundary and our descriptions of the boundary. Every account of this type will do the same. There is the same sort of undamaging circularity in mental state and enjoyment accounts, because they need to get beyond the ordinary senses of 'pleasure' and 'enjoyment', and they would have to go about fixing a new boundary in just the same way.

This narrowing of the desire account still does not get rid of the great embarrassment of the desires of the dead. Of course, a lot of the desires of the dead do count morally, but that is because they affect the living. There is a good case for honouring wishes expressed in wills. Inheritance satisfies the desires of the living to provide for their offspring and encourages saving that benefits society generally. There is a good case, too, for granting rights to the dead—say, to

determine whether their bodies are used for medical purposes.[2] But that, again, does not require appeal beyond the well-being of the living. And, anyway, that a desire of a dead person counts *morally* does not show that it counts towards his well-being.

The real trouble is our counting the fulfilment of aims even if (as it seems we must) we do not require that the fulfilment enter experience. Some of our aims are not fulfilled until we are dead; some, indeed, being desires for then, could not be. But is this so embarrassing, after all?[3] You might have a desire—it could be an informed one, I think—to have your achievements recognized and acknowledged. An enemy of yours might go around slandering you behind your back, successfully persuading everyone that you stole all your ideas, and they, to avoid unpleasantness, pretend in your presence to believe you. If that could make your life less good, then why could it not be made less good by his slandering you with the extra distance behind your back that death brings? You might well be willing to exert yourself, at risk of your life, to prevent these slanders being disseminated after your death. You might, with eyes full open, prefer that course to longer life with a ruined reputation after it. There seems nothing irrational in attaching this value to posthumous reputation. And the value being attached to it does not seem to be moral or aesthetic or any kind other than the value to be attached to the life as a life to be lived.[4] Here is another example. It would not have been at all absurd for Bertrand Russell to have thought that if his work for nuclear disarmament had, after his death, actually reduced the risk of nuclear war, his last years would have been more worthwhile, and his life altogether more valuable, than if it all proved futile. True, if Russell had indeed succeeded, his life clearly would have been more valuable to others. But Russell could also have considered it more valuable from the point of view of his own self-interest. For instance, it would not have been absurd for Russell to think the same about devoting his last years to some purely intellectual project without effects on others' well-being, such as patching up the holes in the Theory of Descriptions. A lot of desires of the dead would be ruled out on the grounds we have already mentioned, but it seems right for some still to count.[5]

2. Why We Should Resist Restricting It More

Excluding some desires raises the general question of whether the best account of 'utility' will not exclude desires of several further sorts. Should not other-regarding desires be excluded?[6] Those who not only want their own welfare but also, luckily for them, have others wanting it too count more heavily than those who do not; for instance, orphans count less than children with loving parents. But that yields Bentham out of Orwell: each to count for at least one but some for more than one. Should not irrational desires be excluded?[7] The principle of utility is a normative principle and ought perhaps, therefore, to grant weight only to what are, by its own standards, good reasons, such as benefit and freedom from harm, and to grant weight only to desires justifiable in terms of these reasons. Should not, for obvious reasons, immoral desires be excluded?[8] Indeed, should not desires of any sort of moral character be excluded?[9] If the concern of the principle of utility is with what ought to be done, then the desire for something because it is what ought to be done appears when the principle delivers its result and seems improper as a ground for the result. Do we not, in general, need a Theory of Types in utility theory to exclude certain desires from the argument place in utility functions?

I think not. First of all, it is impossible to separate self-regarding and other-regarding desires. Each of us wants certain pure states of himself (e.g. to be free from pain); but we also want our lives to have some point, and this desired state can be hard to separate from states of others. Also, if we accepted the restriction to self-regarding desires, we should sever the connection between utility and happiness (and happiness is a large part of utility even on the informed-desire account). A father's happiness can be at stake in his child's happiness—two persons' welfare riding on one person's fate. Allowing that is no violation of everybody's counting for one; it merely allows the father, like everyone else, also to count for one. We have to swallow a little harder when we shift from involvements such as love to involvements such as hate, envy, spite, prejudice, and intolerance. If these sorts of desires are going to count too, what awful distortions will creep into

political decision? But if a lover's happiness counts, so does a hater's *schadenfreude*. It is an ugly sort of pleasure, and as pleasures go slight and troublesomely mixed, but still a pleasure. If it ought not to have weight in moral or political calculation, then we had better find a way to keep it out. But out of moral and political calculation, where it probably does not belong, not out of 'utility', where in some small way it probably does. Also letting other-regarding desires enter moral calculation seems to distort the notion of a moral reason. If one of my aims is to convince Britain that it ought to go over to comprehensive schools, why should you think that my desire constitutes yet another moral reason for going over to comprehensive schools? It would be absurd to introduce comprehensive education because it satisfied the desires of its advocates. What has overwhelming weight here, of course, is the good of the children and of society at large. But all that one really has to swallow is that the happiness of the advocates of comprehensives may in some small way turn on what happens, and that, at least, seems right.[10]

Simply to rule out irrational desires would also go too far. A compulsive hand-washer's desire is irrational, but its fulfilment affects his utility. So since irrational desires cannot be excluded wholesale, why not let them in, and if their fulfilment is sometimes morally intolerable, look to other moral matters besides utility to block it. True, the fulfilment of other sorts of irrational desires is more worrying. A misogynist might be put off his food by a woman's sitting next to him in the Senior Common Room. Consistency would seem to require that his desire not to have women around counts too. Well, why not? The suggestion earlier was that desires that are irrational on utilitarian grounds should not be given weight, because no utilitarian value is at stake. But if someone is upset or distressed, then there is a utilitarian value at stake. The theoretical oddity would come, not in giving weight to such desires, but in giving them none.

What if desires are not only irrational but downright immoral? Should we count, for instance, sadistic desires? This has seemed more of a challenge than it really is, simply because people still tend to think of 'utility' in rather narrow hedonistic terms. Anyone with much understanding would

regard his own sadistic desires—even purely from the point of view of how good his life is for him to lead—as making virtually no claim upon being fulfilled. He would have formed second-order desires not to encourage or indulge them; he would know that, in his case at least, their gratification is mixed and brings no lasting or deep enjoyment; he would know that their opportunity cost is enormous. In fact, it is hard to think of any fairly normal sort who would not be better off, from his own point of view, frustrating his sadistic desires and trying for something better. Still, it would be a mistake simply to rule out sadistic desires. Not everyone is fairly normal. Perhaps there is someone for whom sadistic kicks are all he has, who is incapable of better. It might even be right, if he were also an inept sadist who aims to shock and upset but succeeds only in boring, to play along with him. The same holds for a desire to do something because it would be morally right. There are people for whom living morally is so much at the centre of their lives that their success there is a large part of their lives' being successful. The ideal development of human nature is for 'ethical push' (self-interest) and 'ethical pull' (obligation) to get progressively closer to each other.[11] That too complicates the notion of 'well-being', but desirably so.[12]

All these cases for further restrictions focus not on 'well-being' itself but on how it fits into moral theory. So this is another time when we have to remind ourselves that the question, 'What is the best account of "utility"?', is quite distinct from the question, 'What is the best account, as *ad hoc* as you like, that yields the most adequate one principle, utility maximizing, moral theory?'[13]

3. How Value and Desire Are Related

The danger is that desire accounts get plausible only by, in effect, ceasing to be desire accounts. We had to qualify *desire* with *informed*, and that gave prominence to the features or qualities of the objects of desire, and not to the mere existence of desire. Then, to prevent informed desires from spreading too widely, we had to give prominence to only a certain range of features or qualities. Does this not confirm the suspicion

that desire is no longer playing any real part, and remains only as a token of piety to a utilitarian tradition that has now effectively been abandoned? The issue widens. Could desire be a *ground* of value, or is it at best only a *mark* of it? Are things valuable because desired, or desired because valuable? And widest of all, what place do reason (cognition, perception, judgment) and desire (will, appetite, conation) have in explaining value?[14]

In a way the order of explanation must be from *value* to *desire*.[15] We see that an object has certain features, such as that it is pleasant or healthy or that it gives security, or that it would be an accomplishment. And therefore we desire it. We have always to be able to cite some feature that makes the desirability of the object intelligible; otherwise the notion of 'value' loses hold. And that feature has to be generally intelligible as one that makes things desirable. No one can just make something valuable by adopting it as his own personal aim.[16] Of course, people can disagree in their values. I might find mountain-climbing exciting and value it highly; you may find it simply terrifying and not value it at all. But we do not disagree here in our values in any deep or interesting way. Virtually everyone values excitement and does not value pure terror, though people differ in what they find exciting and terrifying. We all have to be able to connect what we value to some generally intelligible desirability feature. What is more, we sometimes discover values. You may be happy-go-lucky and not even think about accomplishing anything with your life, but then come upon someone whose accomplishment makes his life seem to you exhilarating and fulfilled. And with time you may come to discover more and more what this desirable accomplishment really is; you see how to separate it from mere achievement and its value from merely gaining praise. When you see what accomplishment is, you form a desire. And there need not be any pre-existent background desire (except those of vacuous generality) of which your new desire is merely another instance.

But what is interesting is how little any of this shows. It still leaves a strong case for saying that the order of explanation is quite the reverse: from *desire* to *value*. True, objects are valuable because of *their* features. But how do we explain these

various desirability features? How do we separate *desirability* features from the rest? Here we have to guard against taking one or two examples as paradigms, and missing the variety of cases. So consider the following ones.

Case 1: I have tasted both apples and pears. I like both but prefer pears. How do we explain my attaching more value to having a pear? The only relevant desirability feature is that they taste good. However, it is not a plausible explanation of *tasting better* that I perceive that pears possess this desirability feature to a greater degree than apples. We need to explain my liking pears more in terms of my wanting them more. That is true whether different persons' tastes coincide or not. But another important feature of tastes is that often they do not. We have no reason to expect, with many tastes, that differences in valuing shows that there is any lack of perception or understanding. My preference for pears is not open to criticism (though others of my tastes are—for lack of discrimination, experience, attention). There is a tradition, especially strong in the social sciences, that sees all preference on the model of the simplest tastes: a pre-existent motivation, not subject to criticism, unaffected by understanding; the explanation running from *desire* to *value*. But this is only one kind of case.

Case 2: A recluse may see what he is missing and come to prefer good company. Here perception plays a large part; it may even be a case of discovering a value. But why is good company itself seen as desirable? The explanation cannot be just in terms of perception; there is an important pre-existent motivation. The motivation is not a taste, which typically can vary from person to person; it is more a feature of human nature. We are social creatures; we want and, other things being equal, go for company.[17]

Case 3: Freud, in his last days, preferred thinking clearly to drugged comfort. Here there is a large element of perception but no obvious, at least simple, pre-existent motivation. We get no plausible explanation of this case unless we bring in both understanding what it is to think clearly and wanting it more. Explaining the state of *thinking clearly* as a desirability feature needs both perception and desire, without priority to either. To see this feature as desirable and to desire it on seeing

it are the same. There is no plausible explanation of the one in terms of the other.[18]

Case 4: A person who in the past has frittered his life away comes to value accomplishing something with it. Here understanding plays an enormous role, and desire may seem to disappear altogether. Accomplishment, in the general sense I have in mind (making one's life valuable and not just frittering it away), is valuable for everyone; anyone who fails to recognize it as valuable lacks understanding. It is true that there will be odd types for whom, all things considered, it will be better not to try for it. Perhaps someone for whom any ambition sets up intense anxieties had better not. But that is a case of conflict of values, in which accomplishment is still a value. Does the priority now run the other way? Do we see a life of accomplishment as valuable and then, on the basis of that, form a desire for it? Clearly, this is not a case of first perceiving facts neutrally and then desire's entering and blindly fixing on one object and not another. The way in which we talk about the objects we value is far from neutral; we call it 'accomplishment' and explain it in terms of 'giving life weight and substance' or 'not wasting life'. The language we use in reporting our perceptions already organizes our experience and selects what we see as important; it is designed to show how we view certain things in a favourable light. Desire here does not blindly fix on an object; it is obviously pointed in certain directions by what we perceive favourably. But all of this, though true, explains too little. We also have to explain what goes on in our perceiving things favourably. And here desire comes back at a deeper level, as part of this explanation. Hume was wrong to see desire and understanding (appetite and cognition) as distinct existences. He was wrong to make desire blind. But it is a variety of the same mistake to think that one can explain our fixing on desirability features purely in terms of understanding. It is a mistake not only to keep understanding out of all desire but also to keep desire out of all understanding. Some understanding—the sort that involves fixing on certain features and seeing them in a favourable light—is also a kind of movement.[19] It requires a will to go for what has those features. There is no adequate explanation of their being *desirability* features without appeal

to this kind of movement. So we cannot, even in the case of a desirability feature such as *accomplishment*, separate understanding and desire. Once we see something as 'accomplishment', as 'giving weight and substance to our lives', as 'avoiding wasting our lives', there is no space left for desire to follow along in a secondary, subordinate position. Desire is not blind. Understanding is not bloodless. Neither is the slave of the other. There is no priority.

It may still seem that a value such as accomplishment has to have some priority to desire. It may seem that such a value cannot even depend upon what, if informed, persons would desire. One thing that would show this is our deciding that accomplishment is of absolute value. But few of us believe that; it is much more plausible to think there could be the very rare case in which trying to accomplish something was so painful that it was not worth doing. Another thing that would show it is our deciding that the value of accomplishment, while not absolute, is not given by its place in informed desires either, in other words that reflectively wanting accomplishment and recognizing its value to one can differ. But what would the difference be? Where, for instance, would trade-offs sanctioned by these different conceptions of the value of accomplishment diverge?

So desire is more than merely a mark of value. It is a ground, in the following sense: it is part of the full explanation of prudential value. But this does not give desire priority. Nor does the appearance of *desiring* in *valuing* mean that we are free to make an existential choice of values. The desires that count are not brute and unconstrained; they are informed. It is the strength of the notion of 'informed desire' that it straddles—that is, does not accept any sharp form of—the divide between reason and desire.

The advantages of the informed desire account, therefore, seem to me to be these. It provides the materials needed to encompass the complexity of prudential value. It has the advantages of scope and flexibility over explanations of 'well-being' in terms of desirability features. It has scope, because all prudential values, from objects of simple varying tastes to objects of universal informed agreement, register somewhere in informed preferences. It has flexibility, because

not everyone's well-being is affected in the same way by a certain desirability feature, and we want a notion sensitive to these individual differences. We want to know not only that something is valuable, but how valuable it is, and how valuable to different persons.