

painful or conducive to pain. But pleasure and pain, though they are thus not the *ratio essendi* of good and evil, are the *ratio cognoscendi* thereof. Pleasure is the infallible sign of heightened vitality, pain is the infallible sign of lowered vitality, and these are the only ultimate good and evil. If a man were born with completely clear ideas and completely active emotions he would, according to Spinoza, have no idea of good or evil. For he would never have felt the pleasure of passing to a higher degree of vitality and mental clearness nor the pain of passing to a lower degree of vitality and to a state of greater mental confusion. Yet he would in fact be in the best state in which a human being could be. But the hypothesis in question is one that could not possibly be realised, for we necessarily start in a state of predominantly confused cognition and predominantly passive emotion. There is just one qualification to be made to the above statements. We must remember the distinction between Well-being and Localised Pleasure, and between Depression and Localised Pain. It is only the first members of these two pairs which are infallible signs of heightened and lowered vitality respectively, and therefore of good and evil.

CHAPTER III

Butler

BUTLER'S ethical theories are contained in the *Sermons on Human Nature* which he preached at the Rolls Chapel in London, and in the *Dissertation on the Nature of Virtue* which forms one of the appendices to his famous *Analogy of Religion*. It would be hard to find two writers of such eminence who were so unlike each other as Butler and Spinoza. The writer with whom he has most affinity among those who are treated in this book is Kant, though Hume accepted and emphasised his refutation of psychological egoism. Butler was not, of course, as great a metaphysician as Kant; but he largely made up for this by his clearness and balance. Kant's work is marred by a mania for neat logical classifications and by a strong trace of moral fanaticism; whilst Butler has the solid common-sense and the sweet reasonableness of an English bishop of the eighteenth century. He writes about facts with which we are all acquainted in language which we can all understand; and his work, though it does not pretend to be a complete treatise on ethics, forms one of the best introductions to the subject that exists.

It is necessary to say something at the outset about the ethical and religious tone of the period, because this largely determined the form in which Butler put his arguments. The Christian religion was then going through one of its recurrent phases of dormancy, and has seldom been at a

lower ebb in England. Although it has undergone much more serious attacks since Butler's time, I should say (speaking as an outside observer) that it is far more alive now than then. Religion was in a resting stage, worn out with the theological excitements of the seventeenth century and awaiting the revival which was to take place in the latter part of the eighteenth. Butler says in his preface to the *Analogy*: "It is come to be taken for granted by many persons that Christianity is not so much a subject of inquiry, but that it is now at length discovered to be fictitious. And accordingly they treat it as if in the present age this were an agreed point among all people of discernment; and nothing remained but to set it up as a principal subject of mirth and ridicule, as it were by way of reprisals for its having so long interrupted the pleasures of the world." This would certainly not be an accurate description of the attitude of "people of discernment" at the present time towards religion in general or Christianity in particular. We do indeed meet with such people; but they strike us as quaint and picturesque survivals of the eighteenth-century who are rendered all the more amusing by their obviously sincere conviction that they are daringly advanced thinkers.

It was also fashionable in Butler's time to deny the possibility of disinterested action. This doctrine, which was a speculative principle with Hobbes, has always had a certain vogue. It is not without a certain superficial plausibility, and it has naturally been popular both with vicious persons who wanted a philosophical excuse for their own selfishness and with decent people who felt slightly ashamed of their own virtues and wished to be taken for men of the world. One of Butler's great merits is to have

pointed out clearly and conclusively the ambiguities of language which make it plausible. As a psychological theory it was killed by Butler; but it still flourishes, I believe, among bookmakers and smart young business men whose claim to know the world is based on an intimate acquaintance with the shadier side of it. In Butler's day the theory moved in higher social and intellectual circles, and it had to be treated more seriously than any philosopher would trouble to treat it now. This change is very largely the result of Butler's work; he killed the theory so thoroughly that he sometimes seems to the modern reader to be flogging dead horses. Still, all good fallacies go to America when they die, and rise again as the latest discoveries of the local professors. So it will always be useful to have Butler's refutation at hand.

After these preliminaries we can consider Butler's ethical theory as a whole. His chief merit is as a moral psychologist. He states with great clearness the principles according to which decent people do feel and act and judge, though they could not state these for themselves. And, in the course of this, he refutes certain plausible fallacies which would not have occurred to common-sense, but which unaided common-sense cannot answer when learned men suggest them to it. His fundamental doctrine is that the human mind is an organised system in which different propensities and principles can be distinguished. But it is not enough to enumerate these without saying how they are related to each other. It would not be an adequate description of a watch to say that it consists of a spring, wheels, hands, etc., nor would it be an adequate description of the British Constitution to say that it consists of the King, Lords, and Commons. We do not understand

the nature of a watch until we know that the spring makes the wheels turn, that the balance-wheel controls them, and that the object of the whole arrangement is to record the time. Similarly, we do not understand the British Constitution till we know the precise functions and the mutual relations of the King, the Lords, and the Commons.

Now Butler explicitly compares the nature of man both to a watch and to a constitution. He says that we do not fully understand it till we know what it is for and what are the various functions and relations of the various principles and propensities. According to him none of these is intrinsically evil. Wrong-doing is always the excessive or inappropriate functioning of some principle of action which is right when acting in its due degree and in its proper place. It is like a watch with a spring which is too strong for its balance-wheel, or a constitution in which one of the estates of the realm usurps the functions of another. So the essential thing about man as a moral being is that he is a complex whole of various propensities arranged in a hierarchy. These propensities have a certain right proportion and certain right relations of subordination to each other. But men can go wrong, just as watches and constitutions can do; and so we must distinguish between the actual relative strength of our various propensities and that which they ought to have. The latter may be called their "moral authority". It may well happen that at times a principle of higher moral authority has less psychological strength than one of lower moral authority. If so the man will be likely to act wrongly. The rightness or wrongness of an action, or even of an intention, can be judged only by viewing it in relation to the whole system

in which it is a factor. Thus we judge very differently the same action or intention in a child or a lunatic or a sane grown man. Similarly we do not blame a motor-car for irregularities which would make us regard a watch as worthless. This is because watches and motor-cars are very different systems with very different functions. An actual motor-car must be judged by comparing its behaviour with that of an ideal car, and an actual watch by comparing its behaviour with that of an ideal watch.

It is pretty clear that Butler has hold of a sound and intelligible idea, which is as old as Plato's *Republic*. He chooses to express his theory in the form that virtue consists in acting in accordance with one's nature, and that vice is acting against one's nature. I am not fond of the words "natural" and "unnatural", because they are extremely ambiguous and are commonly used by people to convey a flavour of moral authority to their personal likes and dislikes. Butler fully recognises this; he sees that in one sense nobody can act against his nature. I think it would be better to say that virtue consists in acting in accordance with the *ideal nature* of man, and that vice consists in acting against it. No man's actual nature is the ideal nature of man. But this raises no special difficulty. We can form the conception of a perfect watch, although no real watches are perfect. And science makes great use of such idealised concepts as perfectly straight lines, perfect circles, perfect gases, etc., though it admits that there are no such objects in Nature.

We must now consider how such concepts are reached, so as to see whether the concept of an ideal human nature is likely to be valid or useful. I think that we commonly reach them in two different ways. In forming the concept

of a perfect watch we start with a knowledge of what watches are for. A watch is for telling the time, and a perfect watch would be one that told the time with complete accuracy. Butler often talks as if we could apply this criterion to man, but this does not seem to me to be true. There is no sense in asking what a man is for unless we assume that he has been made by God for a certain purpose. And, even if this were certain, it would not help us; for we do not know what this purpose may be. But there is another way in which we form ideal concepts, and this is illustrated by the concept of a perfect circle or straight line. We see such things as cakes and biscuits and pennies. On reflection we see that they fall into a series—cake, biscuit, penny—in which a certain attribute is more and more fully realised. Finally we form the concept of a perfect circle as the ideal limit to such a series. Thus we can form the concepts of such ideal limits as circles and straight lines by reflecting on imperfect instances arranged in series; and here there is no need to know what the objects are for. Intermediate between the ideal watch and the ideal circle, and more closely analogous to what Butler needs for his purpose, would be the biologist's concept of an ideal horse or rabbit. By comparing and contrasting actual horses, all of which are defective in various respects and to various degrees, we can form the notion of an ideal horse. And, although we recognise that it is an anthropomorphic way of speaking and that we must not take it too literally, we are making a statement which has some kind of correspondence to an important fact when we say that Nature is always striving towards such ideals and always falling short of them to some extent.

There are three things to notice about these ideal limits.

(1) There is generally no lower limit to such series. There is a concept of a perfectly straight line, but there is no concept of a perfectly crooked one. (2) When we have formed the concept of an ideal limit we sometimes find that it is analysable and sometimes that it is not. We can define "circularity", but we cannot define "straightness". Yet we understand just as well what is meant by one as by the other. (3) We could not reach the concepts of these ideal limits unless we had the power of reflecting on series and recognising the characteristic which is more and more adequately, though still imperfectly, realised in the higher members of the series.

Now I think that there is an exact analogy to these three points in forming the concept of an ideal human nature. (1) There is no concept of a perfectly bad man, any more than there is a concept of a perfectly crooked line. (2) If we arrange actual men, including ourselves, in a series, and reflect on it, we can detect a closer and closer approximation to an ideal which is not exactly realised in any of them. But it does not follow that we can analyse and define this ideal completely. I think that Butler would say that we can indicate its general outlines but not its precise details. It certainly involves, as we shall see, the subordination of particular impulses to the more general principles of prudence and benevolence. And it certainly involves the subordination of both these general principles to the supreme principle of conscience. But just how far each impulse would be indulged in the ideal man, and just what compromise he would make between prudence and benevolence when the two conflict, Butler does not tell us. And perhaps it is impossible for anyone to tell us.

This margin of vagueness does not, however, make the concept of an ideal human nature either unintelligible or useless. (3) Butler would say that we could not form this concept at all unless we had the power of reflecting upon actions and characters and comparing them in respect of moral worth. Moral worth is evidently a characteristic of a quite peculiar kind. It is not considered by the other sciences; and so the ideal gases of physics or the ideal circles of geometry may be called "purely positive ideals" and must be contrasted with the ideal human nature which is contemplated by ethics. The power of recognising this peculiar characteristic is one which we plainly do have and do constantly use. It is the cognitive aspect of what Butler calls *Conscience*. With these explanations it seems to me that Butler's conception of an ideal human nature is sound, and that it is true to say that virtue consists in acting in accordance with this nature, and that vice is acting against it.

We can now consider in greater detail how Butler supposes human nature to be constituted. In all men he distinguishes four kinds of propensities or springs of action: (1) There are what he calls "particular passions or affections". These are what we should call impulses to or aversions from particular kinds of objects. Hunger, sexual desire, anger, envy, sympathy, etc., would be examples of these. It is obvious that some of them mainly benefit the agent and that others mainly benefit other people. But we cannot reduce the former to self-love or the latter to benevolence. We shall go more fully into this very important doctrine of Butler's later. (2) There is the general principle of cool self-love. By this Butler means the tendency to seek the maximum happiness for ourselves over

the whole course of our lives. It is essentially a rational calculating principle which leads us to check particular impulses and to co-ordinate them with each other in such a way as to maximise our total happiness in the long run. (3) There is the general principle of benevolence. This, again, is a rational calculating principle, which must be sharply distinguished from a mere impulsive sympathy with people whom we see in distress. It is the principle which makes us try to maximise the general happiness according to a rational scheme and without regard to persons. I think it would be fair to say that the ideal of the Charity Organisation Society is benevolence in Butler's sense. (4) There is the principle of Conscience which is supreme over all the rest in authority. In ideal human nature conscience is supreme over self-love and benevolence; *i.e.*, it determines how far each of these principles is to be carried. Self-love and benevolence in their turn are superior to the particular impulses; *i.e.*, they determine when and to what extent each shall be gratified. In any actual man self-love may overpower conscience and so spread itself at the expense of benevolence. We then get the coolly selfish man. Or benevolence may overpower conscience and exercise itself at the expense of proper prudence. This happens when a man neglects self-culture and all reasonable care for his health and happiness in order to work for the general welfare. Butler holds that both these excesses are wrong. We do not indeed, as a rule, blame the latter as much as the former. But we do blame it to some extent on calm reflection. We blame the imprudently benevolent man less than the coolly selfish man, partly because his fault is an uncommon one, and partly because it may be beneficial to society to have some men who are too benevolent

when there are so many who are not benevolent enough. Butler does not mention this last reason; but I have no doubt that he would have accepted it, since he holds that the faulty behaviour of individuals is often overruled by Providence for the general good.

Particular impulse, again, may be too strong for self-love or for benevolence or for both. *E.g.*, revenge often leads people to actions which are inconsistent with both benevolence and self-love, and ill-regulated sympathy may have the same effect. In the latter case we have the man who gives excessively to undeserving cases which happen to move his emotions, and who equally violates prudence by the extent of his gifts and benevolence by his neglect of more deserving but less spectacular cases. Butler makes the profoundly true remark that there is far too little self-love in the world; what we need is not *less self-love* but *more benevolence*. Self-love is continually overcome by particular impulses like pride, envy, anger, etc., and this is disastrous both to the happiness of the individual and to the welfare of society at large. Self-love is not indeed an adequate principle of action. But it is at least rational and coherent so far as it goes; and, if people really acted on it consistently, taking due account of the pleasures of sympathy and gratitude, and weighing them against those of pride, anger, and lust, their external actions would not differ greatly from those which benevolence would dictate. This seems to me to be perfectly true. Those actions which are most disastrous to others are nearly always such as no person who was clear-sightedly aiming at the maximum amount of happiness for himself would dream of doing. We have an almost perfect example of Butler's contention in the action of France towards Germany since the war of

1914 to 1918. It has been admirably adapted to producing the maximum inconvenience for both parties, and, if the French had acted simply from enlightened self-interest instead of malice and blind fear, they and all other nations would now be far better off.

The ideal human nature, then, consists of particular impulses duly subordinated to self-love and benevolence, and of these general principles in turn duly subordinated to the supreme principle of conscience. This seems to me to be perfectly correct so far as it goes; and I will now consider in rather more detail each of these constituents of human nature.

1. *Particular Impulses*.—Butler's first task is to show that these cannot be reduced to self-love, as many people have thought before and since his time. It is easy to see that he is right. The object of self-love is one's own maximum happiness over the whole course of one's life. The object of hunger is food; the object of revenge is to give pain to someone who we think has injured us; the object of sympathy is to give another man pleasure. Each of these particular impulses has its own particular object, whilst self-love has a general object, viz., one's own maximum happiness. Again, these particular impulses often conflict with self-love, and this is equally true of those which we are inclined to praise and those which we are inclined to blame. Nor is this simply a question of intellectual mistakes about what will make us happy. A man under the influence of a strong particular impulse, such as rage or parental affection, will often do things which he knows at the time to be imprudent.

In a footnote Butler takes as an example Hobbes's definition of "pity" as "fear felt for oneself at the sight

of another's distress". His refutation is so short and so annihilating that I will give the substance of it as a model of philosophical reasoning. He points out (a) that, on this definition, a sympathetic man is *ipso facto* a man who is nervous about his own safety, and the more sympathetic he is the more cowardly he will be. This is obviously contrary to fact. (b) We admire people for being sympathetic to distress; we have not the least tendency to admire them for being nervously anxious about their own safety. If Hobbes were right admiration for sympathy would involve admiration for timidity. (c) Hobbes mentions the fact that we tend specially to sympathise with the troubles of our friends, and he tries to account for it. But, on Hobbes's definition, this would mean that we feel particularly nervous for ourselves when we see a friend in distress. Now, in the first place, it may be doubted whether we do feel any more nervous for ourselves when we see a friend in distress than when we see a stranger in the same situation. On the other hand, it is quite certain that we do feel more sympathy for the distress of a friend than for that of a stranger. Hence it is impossible that sympathy can be what Hobbes says that it is. Butler himself holds that when we see a man in distress our state of mind may be a mixture of three states. One is genuine sympathy, *i.e.*, a direct impulse to relieve his pain. Another is thankfulness at the contrast between our good fortune and his ill luck. A third is the feeling of anxiety about our own future described by Hobbes. These three may be present in varying proportions, and some of them may be wholly absent in a particular case. But it is only the first that any plain man means by "sympathy" or "pity". Butler makes a very true observation about this theory of Hobbes. He

says that it is the kind of mistake which no one but a philosopher would make. Hobbes has a general philosophical theory that all action must necessarily be selfish; and so he has to force sympathy, which is an apparent exception, into accord with this theory. He thus comes into open conflict with common-sense. But, although common-sense here happens to be right and the philosopher to be wrong, I should say that this is no reason to prefer common-sense to philosophy. Common-sense would *feel* that Hobbes is wrong, but it would be quite unable to say *why* he is wrong. It would have to content itself with calling him names. The only cure for bad philosophy is better philosophy; a mere return to common-sense is no remedy.

We can now leave Hobbes to his fate, and return to the general question of the relation of our particular impulses to self-love. Why should it seem plausible to reduce particular impulses, like hunger and revenge and sympathy, to self-love? The plausibility arises, as Butler points out, from two confusions. (i) We confuse the ownership of an impulse with its object. All our impulses, no matter what their objects may be, are *ours*. They all *belong to* the self. This is as true of sympathy, which is directed to others, as of hunger, which is directed to modifying a state of oneself. (ii) Again, the satisfaction of any impulse is *my* satisfaction. *I* get the pleasure of satisfied desire equally whether the desire which I indulge be covetousness or malice or pity. So it is true that all impulses *belong to* a self, and that the carrying out of any impulse as such *gives pleasure to* that self. But it is not true that all impulses have for their objects states of the self whose impulses they are. And it is not true that the object of any of them is the general happiness of the self who owns them. Neither sympathy

nor malice is directed to producing the happiness of the self who owns these impulses. One is directed to producing happiness in another person, and the other is directed to producing misery in another person. Thus there is no essential contrariety between any impulse and self-love. The satisfaction of any of my impulses as such gives me pleasure, and this is a factor in that total happiness of myself at which self-love aims. And self-love can gain its end only by allowing the various special impulses to pass into action. On the other hand, no impulse can be identified with self-love. The relation of particular impulses to self-love is that of means to end, or of raw materials to finished product.

All this is true and very important. But to make it quite satisfactory it is necessary, I think, to draw some distinctions which Butler does not. (i) We must distinguish between those pleasures which consist in the fulfilment of pre-existing desires and those which do not. Certain sensations are intrinsically pleasant, *e.g.*, the smell of violets or the taste of sugar. Others are intrinsically unpleasant, *e.g.*, the smell of sulphuretted hydrogen or the feel of a burn. We must therefore distinguish between intrinsic pleasures and pains and the pleasures and pains of satisfied or frustrated impulse. All fulfilment of impulse is pleasant for the moment at least; and all prolonged frustration of impulse is unpleasant. This kind of pleasure and pain is quite independent of the object of the impulse. Now these two kinds of pleasure and pain can be combined in various ways. Suppose I am hungry and eat some specially nice food. I have then both the intrinsically pleasant sensation of taste and also the pleasure of satisfying my hunger. A shipwrecked sailor who found some putrid meat or dined

off the cabin-boy would enjoy the pleasure of satisfying his hunger accompanied by intrinsically unpleasant sensations of taste. A *bon-vivant* towards the end of a long dinner might get an intrinsically pleasant sensation of taste from his savoury although he was no longer hungry and therefore did not get the pleasures of satisfying his hunger.

(ii) I think that we must distinguish between the object of an impulse, its exciting cause, what will in fact satisfy it, and the collateral effects of satisfying it. Butler lumps together hunger and sympathy, and says that the object of one is food and the object of the other is the distresses of our fellow-men. Now, in the first place, the word "hunger" is ambiguous. It may mean certain organic sensations which are generally caused by lack of food. Or it may mean an impulse to eat which generally accompanies these. Butler evidently uses the word in the latter sense. But, even in this sense, it seems to me inaccurate to say that the object of hunger is food. It would be equally true to say that the object of a butcher going to market is food; but he may not be hungry. The object or aim of hunger is *to eat food*. The object of the butcher is to buy it as cheaply and sell it as dearly as possible. In fact the object of an impulse is never, strictly speaking, a thing or person; it is always to change or to preserve some state of a thing or person. So much for the object or aim of an impulse.

Now, as we eat, the impulse of hunger is gradually satisfied, and this is pleasant. If we are continually prevented from eating when we are hungry this continued frustration of the impulse is unpleasant. Lastly, the process of satisfying our hunger has the collateral effect of producing sensations of taste which may be intrinsically pleasant

or unpleasant according to the nature of the food and the tastes of the eater. I would say then that the exciting cause of the impulse of hunger is lack of food, accompanied in general by certain characteristic organic sensations; that its aim or object is the eating of food; that its collateral effects are sensations of taste; and that it is accompanied by satisfaction or dissatisfaction according to whether we get food or are unable to do so. Now let us consider pity from the same points of view. The exciting cause is the sight of another person, particularly a friend or relation, in distress. The aim or object of it is to relieve the distress. The collateral effects of its exercise are the gradual relief of the distress, feelings of gratitude in the sufferer's mind, and so on. Lastly, in so far as we are able to exercise the impulse, there is a pleasant feeling of satisfaction in our minds; and, in so far as we are prevented from doing so, there is an unpleasant feeling of frustration.

Now, in considering the relations between the various particular impulses and the general principles of self-love and benevolence, it is very important to bear these distinctions in mind. Butler says that some particular impulses are more closely connected with self-love and others with benevolence. He gives examples, but he does not carry the analysis further. We can now state the whole case much more fully and clearly. (a) Some impulses have their exciting causes in the agent, some in inanimate objects, and some in other persons. Hunger is excited by one's own lack of food and the organic sensations which accompany it; covetousness may be excited by the sight of a book or a picture; pity is excited by another man's distress. (b) Some impulses aim at producing results within the agent himself; some aim at producing results in other

men; and some aim at effecting changes in inanimate objects. Thus hunger aims at one's own eating; pity aims at the relief of another man's distress; and blind rage may aim at smashing plates or furniture. (c) The collateral effects of satisfying an impulse may be in the agent, or in others, or in both. Probably there are always collateral effects in the agent himself, and nearly always in other men too. But sometimes the collateral effects in the agent predominate, and sometimes those produced in other men are much more important. The collateral effects of satisfying hunger are, under ordinary circumstances, almost wholly confined to the agent. The collateral effects of the exercise of pity are mostly in the sufferer and the spectators, though there are always some in the agent. The collateral effects of ambition are divided pretty equally between self and others. Lastly, (d), the pleasures of satisfied impulse and the pains of frustrated impulse are naturally confined to the owner of the impulse.

It is evident that those particular impulses which aim at producing or maintaining states of the agent himself, and those whose collateral effects are mainly confined to the agent, will be of most interest to self-love. Hunger is a typical example. Those impulses which aim at producing or altering or maintaining states in other men, and whose collateral effects are mainly confined to others, will be of most interest to benevolence. Sympathy and resentment are typical examples. There will be some impulses which almost equally concern self-love and benevolence. For it may be that they aim at producing a certain state in others, but that their collateral effects are mainly in the agent; or conversely. Anger against those whom we cannot hurt is aimed against them but mainly affects ourselves. The

of some particular person, *e.g.*, paternal and filial affection. He says that, if you grant that paternal and filial affection exist, you must grant that benevolence exists. This is a mistake. He might as well say that, if you grant that hunger exists, you must grant that self-love exists. Really paternal affection is just as much a particular impulse as hunger, and it can no more be identified with benevolence than hunger can be identified with self-love. I think that he makes such apparent mistakes partly because he is anxious to show that benevolence is, as such, no more contrary to self-love than is any of the particular impulses. He shows, *e.g.*, that to gratify the principle of benevolence gives just as much pleasure to the agent as to gratify any particular impulse, such as hunger or revenge. It is true that excessive indulgence in benevolence may conflict with self-love; but so, as he points out, may excessive indulgence of any particular impulse, such as thirst or anger. In fact benevolence is related to self-love in exactly the same way as any particular impulse is related to self-love. So far he is perfectly right. But this identity of relation seems sometimes to blind Butler to the intrinsic difference between benevolence, which is a general principle, and the particular impulses which aim at producing happiness in certain particular men or classes of men, *e.g.*, patriotism or paternal affection.

I think that there is undoubtedly a general principle of benevolence; and I think that Butler held this too, though he does not always make this clear. The main business of benevolence is to control and organise those impulses which aim at producing changes in others, or whose collateral effects are mainly in others. Thus it has to do with pity, resentment, paternal affection, and so on. The main business

of self-love is to control and organise those impulses which aim at producing states in oneself, or whose collateral effects are largely in oneself. From the point of view of self-love benevolence is simply one impulse among others, like hunger, resentment, etc. But it is equally true that, from the point of view of benevolence, self-love is only one impulse among others. The prudent person may need to check his excessive benevolence towards mankind in general, just as he has to check blind anger or a tendency to over-eating. The benevolent person may need to check his excessive prudence, just as he has to check the special impulse to lose his temper.

There are, however, two respects in which self-love and benevolence seem to me to be not perfectly on a level. Conscience approves both of self-love and of benevolence in their proper degrees. But I think it is clear that conscience rates benevolence higher than self-love. It would hold that it is possible, though not easy, to have too much benevolence, but that you could quite easily have too much self-love, though in fact most people have too little. Again, from a purely psychological point of view, self-love and benevolence are not quite co-ordinate. The putting into action of *any* tendency, including benevolence, is as such pleasant to the agent, and so ministers in its degree to self-love. But the putting into action of our conative tendencies is not as such a source of happiness to others. Others may be affected either pleasurably or painfully according to the nature of the impulse which I exercise. But I get a certain amount of pleasure from the mere fact that I am doing what I want to do, quite apart from whether the object of the action is my own happiness or whether its collateral consequences are pleasant sensations in myself. Thus no action of mine

can be *completely* hostile to self-love, though the collateral results of the action may be so unpleasant for me that cool self-love would not on the whole sanction it. But the gratification of many impulses may be completely hostile to benevolence. If I lose my temper and blindly strike a man, self-love gets something out of the transaction, viz., the momentary feeling of satisfaction at fulfilling an impulse, even though the remoter consequences may be so unpleasant for me that cool self-love would have prevented the action. But benevolence gets nothing out of the transaction at all; it is wholly hostile to it.

As we have said, Butler holds that pure self-love and pure benevolence would lead to very much the same external actions, because the collateral results of most actions really make about as much for the happiness of the agent as for that of others. In this connexion he makes two profoundly true and important observations. (i) If you want to make yourself as happy as possible it is fatal to keep this object constantly before your mind. The happiest people are those who are pretty fully occupied with some activity which they feel to be honourable and useful and which they perform with reasonable success. The most wretched lives are led by men who have nothing to do but think of their own happiness and scheme for it. Happiness which is deliberately sought generally turns out to be disappointing, and the self-conscious egoist divides his time between wanting what he has not and not wanting what he has. (ii) The second point which Butler makes is that the common opinion that there is an inevitable conflict between self-love and benevolence is a fallacy based on the common confusion between enjoyment itself and the means of enjoyment. If I have a certain sum of money, it is

evident that the more I spend on myself the less I shall have to spend on others, and conversely. It therefore looks at first sight as if self-love and benevolence must necessarily conflict. But, as Butler says, money and other kinds of property are not themselves happiness; they are only material objects which produce happiness by being used in certain ways. Now it is certain that both spending money on myself and spending it on others may give me happiness. If I already spend a good deal on myself it is quite likely that I shall gain more happiness by spending some of it on others than I shall lose by spending that much less on myself. This is certainly true; and the confusion between happiness and the means to happiness, which Butler here explains, is constantly made. The miser illustrates the typical and exaggerated form of this mistake; but nearly every one makes it to some extent.

I think there is only one point in Butler's theory of the substantial identity of the conduct dictated by self-love and by benevolence which needs criticism. It assumes an isolated purely selfish man in a society of people who are ruled by benevolence as well as by self-love and who have organised their social life accordingly. In such a case it certainly would pay this individual to act very much as the principle of benevolence would dictate. It is not so clear that it would pay to act in this way in a community of men who were all quite devoid of benevolence. All that we can say is that every one in such a society, if it could exist at all, would probably be very miserable; but whether one of them would be rendered less miserable by performing externally benevolent actions it is difficult to say. But, if we suppose Butler to mean that, taking men as they are, and taking the institutions which such men have made for

themselves, enlightened self-interest would dictate a line of conduct not very different from that which benevolence would dictate, he seems to be right.

This fact, of course, makes it always difficult to say how far any particular action has been due to benevolence and how far to self-love. What is certain is that both principles exist, and that very few actions are due to one without any admixture of the other. Sometimes we can see pretty clearly which principle has predominated, but this is as far as we can safely go. Exactly the same difficulty arises as Butler points out, over self-love and the particular impulses. It is often impossible to say whether a certain course of action was due to self-love or to a particular impulse for power or money. All that we know for certain is that both principles exist and that they mix in all proportions. Sometimes the onlookers can tell more accurately than the agent what principle predominated, because they are less likely to be biased.

3. *Conscience*.—We come now to Butler's supreme principle of conscience. According to him this has two aspects, a purely cognitive and an authoritative. In addition, I think we must say that it is an active principle; *i.e.*, that it really does cause, check, and modify actions. In its cognitive aspect it is a principle of reflection. Its subject-matter is the actions, characters, and intentions of men. But it reflects on these from a particular point of view. In one sense we are reflecting on our actions when we merely recall them in memory and note that some turned out fortunately and others unfortunately. But we should not call such reflection an act of conscience, but only an act of retrospection. The peculiarity of conscience is that it

reflects on actions from the point of view of their rightness or wrongness. The very fact that we use words like "right", "wrong", "duty", etc., shows that there is an intellectual faculty within us which recognises the terms denoted by these names. Otherwise such words would be as meaningless to us as the words "black" and "white" to a man born blind. We clearly distinguish between a right action and one that happened to turn out fortunately. And we clearly distinguish between a wrong action and one that happened to turn out unfortunately. Again, we distinguish between mere unintentional hurt and deliberate injury. Conscience is indifferent to the former and condemns the latter. Finally, conscience recognises a certain appropriateness between wrong-doing and pain and between right-doing and happiness; *i.e.*, it recognises the fact of merit or desert. If we see a man being hurt we judge the situation quite differently according to whether we think that he is innocent or that he is being punished for some wrong act.

So we may say that conscience, on its cognitive side, is a faculty which reflects on characters, actions, and intentions, with a special view to their goodness or badness, rightness or wrongness. And it further judges that pain is appropriate to wrong-doing, and happiness to right-doing. Lastly, we must add that it does not judge of actions or intentions in isolation, but judges them in reference to the ideal nature of the agent. The ideal nature of a child or a lunatic is different from that of a sane grown man, and so conscience takes a different view of the same action when performed by one or the other. Butler apparently assumes that, although the ideal nature of a child or a lunatic is different from that of a sane grown man, the ideal nature of all mature men is identical. No doubt we

have to assume this in practice ; but it seems hardly likely to be strictly true. It is hard to draw a perfectly sharp line between maturity and immaturity, or between sanity and insanity.

By saying that conscience has supreme authority Butler means that we regard the pronouncements of conscience, not simply as interesting or uninteresting statements of fact, and not simply as reasons to be balanced against others, but as *conclusive* reasons for or against doing the actions about which it pronounces. The fact that conscience pronounces an act to be wrong is admittedly one motive against doing it. But so too is the fact that self-love condemns it as imprudent, or that benevolence condemns it as likely to diminish the general happiness. Thus far conscience, self-love, and benevolence are all on a level. They are all capable of providing motives for acting or abstaining from action. The difference lies in their respective *authority*, i.e., in the relative strength which they *ought* to have and which they *would* have in an ideal human being. If self-love and benevolence conflict over some proposed course of action there is nothing in the nature of either which gives it authority over the other. Sometimes it will be right for self-love to give way to benevolence, and sometimes it will be right for benevolence to give way to self-love. But conscience is not in this position. In an ideal man conscience would not simply take turns with benevolence and self-love. If benevolence or self-love conflict with conscience it is always they, and never it, which should give way ; and, if they conflict with each other, it is conscience, and it alone, which has the right to decide between them. In any actual man conscience is often overpowered by self-love or benevolence, just as they are

often overpowered by particular impulses. But we recognise the *moral right* of conscience to be supreme, even when we find that it lacks the necessary *psychological power*.

I do not think that Butler means to say that every trivial detail of our lives must be solemnly debated before the tribunal of conscience. Just as the man whose aim is to secure his own maximum happiness best secures that end by not constantly thinking about it, so I should say that the man who wants always to act conscientiously will often do best by not making this his explicit motive. So long as our actions are those which conscience would approve, if we carefully considered the question, the supremacy of conscience is preserved, even though we have acted from immediate impulse or self-love or benevolence. Conscience, e.g., approves of a due measure of parental affection ; but it is much better for this affection to be felt spontaneously than to be imposed on a parent by conscience as a duty. In fact the main function of conscience is regulative. The materials both of good and of evil are supplied by the particular impulses. These are organised in the first instance by self-love and benevolence, and these in turn are coordinated and regulated by conscience. In a well-bred and well-trained man a great deal of this organisation has become habitual, and in ninety-nine cases out of a hundred he does the right things without having to think whether or why they are right. It is only in the hundredth specially perplexing or specially alluring situation that an explicit appeal to conscience has to be made.

It remains to say something about two rather curious and difficult points in Butler's theory. (1) Although he constantly asserts the supremacy of conscience, yet there are one or two passages in which he seems to make self-

love co-ordinate with it. In one place he actually says that no action is consistent with human nature if it violates *either* reasonable self-love or conscience. In another famous passage he seems to admit that, if we reflect coolly, we can justify no course of action which will be contrary to our happiness. The former passage I cannot explain away; it seems to be simply an inconsistency. But the latter occurs in the course of an argument in which he is trying to prove to an objector that there is no real conflict between conscience and enlightened self-love. I think it is clear from the context that he is not here asserting his own view, but is simply making a hypothetical concession to an imaginary opponent. He goes on to argue thus. Even if you grant that it can never be right to go against your own greatest happiness, yet you ought to obey conscience in cases of apparent conflict between it and self-love. For it is very difficult to tell what will make for your own greatest happiness even in this life, and it is always possible that there is another life after this. On the other hand, the dictates of conscience are often quite clear. Thus we can be far more certain about what is right than what is to our own ultimate interest; and therefore, in an apparent conflict between the two, conscience should be followed since we cannot be sure that this is not really to our own interest.

So Butler would probably answer that the question whether conscience is superior to self-love or co-ordinate with it is of merely academic interest. I do not think that this answer can be accepted. In the first place, as moralists we want to know what *should be* the relative positions of conscience and self-love. And it is no answer to this question to say that it is not practically important. Secondly, we may grant all that Butler says about the extreme un-

certainty as to what is to our own ultimate interest. But the deliveries of conscience are by no means so certain and unambiguous in most cases as Butler makes out. And even if they were, it is not obvious why they should be assumed to be likely to be a better guide to our own interest than the best opinion that we could reach by reflecting directly on that subject.

(a) The other doubtful point is Butler's view about the value of happiness. In one place he says that it is manifest that nothing can be of consequence to mankind or to any creature but happiness. And he goes on to assert that all common virtues and vices can be traced up to benevolence and the lack of it. Finally, in the same sermon he says that benevolence seems in the strictest sense to include all that is good and worthy. Now, if these statements be accepted at their face-value, Butler was a Utilitarian; *i.e.*, he thought that happiness is the only intrinsic good and that virtue consists in promoting happiness. But it is to be noted that these remarks occur in the sermon on the *Love of our Neighbour*, where he is specially concerned to recommend benevolence to people who were sadly lacking in it. And even here he adds a footnote in which he distinctly says that there are certain actions and dispositions which are approved altogether apart from their probable effect on general happiness. He asserts this still more strongly in the *Dissertation on Virtue*, which is a later and more formal work. So I think it is clear that his considered opinion is against Utilitarianism.

But in both these works he seems to take the interesting view that God may be a Utilitarian, though this is no reason for our being so. It may be that God's sole ultimate motive is to maximise the total amount of happiness in the

universe. But, even if this be the only thing of which he approves as an end, he has so made us that we directly approve of other tendencies beside benevolence, *e.g.*, justice and truth-telling. And he has provided us with the faculty of conscience, which tells us that it is our duty to act in accordance with these principles no matter whether such action seems to us likely to increase the general happiness or not. It is quite possible that God may have given us this direct approval of truth-telling and justice, not because he directly approves of them, but because he knows that it will in fact make for the greatest happiness on the whole if we act justly and speak the truth regardless of the apparent consequences to ourselves and others. If so, that is his business and not ours. Our business is to act in accordance with our consciences, and only to promote the general happiness by such means as conscience approves, even though we may think that we could promote it more in certain cases by lying or partiality. If God does overrule our conscientious actions in such a way that they do make for the greatest possible happiness even when they seem to us unlikely to have that effect, so much the better. It makes no difference to our duty whether this be so or not.

It is of course plain that Butler leaves undiscussed many questions with which any complete treatise on Ethics ought to deal. We should like to know whether there is any feature common and peculiar to right actions, which we could use as a criterion of rightness and wrongness. And we should like to know how, when the same conscience at different times, or different consciences at the same time, seem to issue conflicting orders, we are to tell which is genuine and which is spurious. To such questions Butler

does not attempt to give an answer, whilst the Utilitarians on the one hand and Kant on the other do give their respective very different answers to it. But, though his system is incomplete, it does seem to contain the prolegomena to any system of ethics that can claim to do justice to the facts of moral experience.