CHAPTER IV

Hume

THE best account of Hume's theory of ethics is to be found in his *Enquiry Concerning the Principles of Morals*. This is a treatise specially devoted to ethics. Spinoza's ethical theory is only a part, though a vitally important part, of an elaborate metaphysical theory of the universe. Hume had no such system, and believed it to be impossible for human beings to build one capable of standing. Still, he had certain very definite epistemological principles or prejudices, and these inevitably determined and coloured his ethical theories. The two men were in many ways extremely unlike each other in disposition, outlook, training, and experience; and the spirit of their respective philosophical writings is profoundly different. Yet, in spite of real and important disagreements, we shall find several points of fundamental similarity between the ethical views of Hume and Spinoza.

It will be best, in the case of Hume, to take first that part of ethical theory which we took last in the case of Spinoza, viz., the question of the meaning and analysis of ethical predicates and propositions. Hume's doctrine is the following. There is a certain specific kind of emotion which nearly all human beings feel from time to time. This is the emotion of *Approval* or *Disapproval*. It is called forth by the contemplation of certain objects, and it is directed towards those objects. Now for Hume the

For Hume 'x is good' means that if I were to contemplate x I would experience an emotion of approval.

this entails that good & evil are relational.

it is also a kind of psychological theory.

statement "x is good" *means* the same as the statement "x is such that the contemplation of it would call forth an emotion of approval towards it in all or most men". The definition of "x is bad" would be the same with "disapproval" substituted for "approval".

The following points may be noticed at once. (1) It makes "good" and "bad" to be relational predicates. Their very meaning involves a relation to the human species. So far it resembles Spinoza's view. (2) It is a psychological theory, since it defines "good" and "bad" by reference to certain kinds of mental state, viz., certain kinds of emotion. In this it differs from Spinoza's view. "Good" and "bad", for him, were definable in terms of specific powers and activities. No reference to emotion entered into the *definition*, though he held that the feelings of pleasure and pain are trustworthy *signs* of the presence of good and evil respectively. (3) Though Hume's theory is relational and psychological, in the senses explained, it is not subjective in the sense that it leaves no room for argument and refutation in ethical matters. It would be so if it asserted that "x is good" means "I here and now have an emotion of approval towards x". Such statements, if false, could hardly be refuted; and all argument about them would be unprofitable. But Hume's theory is that "x is good" means that the contemplation of x will call forth an emotion of approval in all or most men on all or most occasions. Such statements as this can be argued about and supported or refuted by observation and collection of statistics. On Hume's theory a man might quite well make the judgment that x is good, though the contemplation of x evoked in him at the time no emotion at all or an emotion of disapproval. For he might acknowledge that

x evokes in most men at most times when they contemplate it an emotion of approval. I think that it is even possible on Hume's theory for a man *first* to judge that x is good, and then, *in consequence of* this judgment, to begin to feel approval of x. For most of us like to feel the same kind of emotions in given circumstances as others feel, especially if we respect or admire the others. And so the mere fact that I believe that most people have a feeling of approval in contemplating x may cause me to feel an emotion of approval in contemplating x which I should not otherwise have felt. (4) I have laboured these points because it is important to see at the outset that such a theory as Hume's does not inevitably lead to such extreme paradoxes that we can reject it out of hand. But we must not underrate the extent to which Hume's theory conflicts with ordinary views. The common view, though it is never very articulately expressed, is presumably somewhat as follows. Certain things would be good and others would be bad whether the contemplation of them did or did not call forth emotions of approval or disapproval in all or most men. The good things call forth emotions of approval in all or most men *because* they are good and *because* men are so constituted as to feel this kind of emotion towards what they believe to be good. And the same is true, *mutatis mutandis*, of bad things. On Hume's view if men did not feel these emotions nothing would *be* good or bad; and it is only in the rather exceptional case which I mentioned above that the judgment that x is good might precede and produce in a certain man an emotion of approval towards x.

Hume now passes to the second part of ethical theory, viz., the question: What kinds of thing are good, and what kinds are bad? This reduces for him to the question:

Is there any characteristic common and peculiar to the things towards which all or most men feel an emotion of approval, beside the fact that they are the objects of this emotion? Hume holds that such a question can be settled only by ordinary observation followed by an empirical generalisation. The result of his observation is that actions, qualities, and characters which are generally approved by men all fall into two classes, viz.: (1) those which are immediately pleasant either to their possessor or to other men; and (2) those which are useful, *i.e.*, ultimately and indirectly productive of pleasure, either to their possessor or to other men. Of course these classes are not mutually exclusive. A benevolent act may be directly pleasant to the agent and to spectators whilst it is useful to the person for whose benefit it is done. And an industrious character is useful both to its possessor and to society. Hume also finds that the converse proposition holds; *i.e.*, everything that falls into one of these classes calls forth an emotion of approval in all or most men who contemplate it. He now generalises these observations by problematic induction, and reaches the conclusion that *all* things which are either directly pleasant or indirectly conducive to pleasure, whether in their owners or in other men, evoke the emotion of approval in all or most men; and that *only* such things do so.

I will now make some comments on this doctrine. (1) In the first place there are two slight ambiguities to be noticed and removed. The first concerns the distinction between what is immediately pleasant and what is useful. There is an ambiguity in the word "pleasant", which may be brought out in the following way. We should commonly say both that chocolate is pleasant and that the experience of tasting chocolate is pleasant. But we should not call

chocolate itself "a pleasure", whilst we should call the experience of tasting it "a pleasure". A pleasure is always a mental event, such as a feeling; or a whole of which a mental event is an essential constituent, though something non-mental may be contained as object, such as hearing a tune, tasting chocolate, etc. Now the word "pleasant" has a different meaning when applied to an experience and to a non-experience; and the former meaning is the fundamental one. In the first sense it denotes a *non-causal* characteristic; in the second it denotes a *causal* characteristic, *i.e.*, a more or less permanent tendency to produce, in co-operation with other factors, a result of a certain kind. Thus to say that a certain tune is pleasant means that it is such that the experience of hearing it will at most times and in most men be pleasant in the non-causal sense, *i.e.*, will be a pleasure. It must be noted that the same thing may be cognised in several different ways; *e.g.*, we can see a bit of chocolate instead of tasting it, we can feel a picture instead of seeing it, and so on. Now it will often happen that some of these different modes of cognising a given object are pleasant experiences whilst the others are neutral. But I think that we call an object pleasant if there be *any* way of cognising it which is a pleasant experience to most men at most times.

I can now define the statement that x is "immediately pleasant". It means that x is either (*a*) a pleasant experience, or (*b*) is such that there is at least one mode of cognising it which is for most men and at most times a pleasant experience. We can now deal with the statement that x is "useful". A thing is useful without being pleasant when it is not itself a pleasure, and when no mode of cognising it is a pleasure, but when it is a cause-factor in the production of

pleasures. It is of course quite possible that one and the same event should be non-causally pleasant, causally pleasant, and useful. Most pleasant experiences are causally pleasant too, since the introspective contemplation of one's own pleasures is itself as a rule a pleasure. And no doubt they are also often cause-factors tending to produce other pleasant experiences in the future, and are thus useful.

The second ambiguity is this. Ought we not to substitute "believed to be" for "are" in Hume's generalisation? Ought we not to say that the emotion of approval is called forth by all those things and only those things which *are believed* by the observer to be immediately pleasant or useful? Presumably things would call forth this emotion if they were believed to have the property, even though they did not in fact have it; and presumably they would not call forth the emotion if they were believed not to have the property, even though they in fact had it. On the other hand, the term "belief" must be taken rather widely if we are not to fall into an opposite error. It must be taken to include what I should call "quasi-belief"; *i.e.*, cases in which we are not explicitly believing or disbelieving so-and-so, but are acting *as if* we believed it, and, if challenged, would explicitly believe it. I do not think that Hume would have objected to either of these modifications in his doctrine; and I shall henceforth assume that they have been made.

(2) My second comment is this. If "Hedonism" be defined as the theory that there is a universal and reciprocal connexion between goodness and pleasantness, then Hume is a hedonist. For he has asserted that everything that is good, in his sense, is pleasant or conducive to pleasure; and that everything which is pleasant or conducive to

pleasure is, in his sense, good. But there are three fundamentally different possible types of hedonism, and Hume's is perhaps the least usual kind. We may first divide hedonistic theories into *Analytic* and *Synthetic Hedonism.* Analytic hedonism asserts that to be "good" *means* to be "pleasant or conducive to pleasure". This is plainly not Hume's view. He is then a synthetic hedonist. But synthetic hedonism may take two forms, *a priori* and *empirical.* The *a priori* synthetic hedonist, whilst denying that "good" means "pleasant or conducive to pleasure", holds that he can see a necessary and reciprocal connexion between the two characteristics, such as we can see between the two characteristics of being equilateral and being equiangular in the case of a triangle. Anything that was good *would necessarily be* pleasant or conducive to pleasure, and conversely. This is the view of such a hedonist as Sidgwick; but it is plainly not Hume's view. The connexion for him is contingent, and the evidence for it is observation and empirical generalisation thereof. He is thus an empirical hedonist. It is logically possible that all or most men should have been so constituted as to feel approval when they contemplated what is painful or conducive to pain in human beings. If so, character and conduct of this kind would have been good. Or, again, men might have been so constituted that they simply did not have the emotions of approval or disapproval at all. If so, nothing would have been either good or bad.

It is then, according to Hume, an empirical and contingent fact that men are so constituted as to feel approval and disapproval, and that they are so constituted that their approvals and disapprovals take the particular direction which he has found that they do take. I propose to call

the innate disposition to feel emotions of approval and disapproval from time to time the *Moral Sentiment.* In order to account for the particular direction which these emotions take in human beings Hume holds that it is necessary to postulate the existence in them of another sentiment, which he calls that of *Benevolence* or *Humanity.* Men are so constituted that every man tends to feel pleased when he contemplates the happiness of any human being and tends to feel displeased when he thinks of any human being as unhappy. There are four points to notice about this emotional disposition. (i) It is common to all, or nearly all, men, like the sexual instinct. (ii) It is excited by the perception or the thought of *any* human being, as such, in a state of happiness or misery. It thus differs, *e.g.*, from self-love or patriotic sentiment. These are no doubt common to most men; but the object which evokes them is a certain man or a certain restricted class of men, not any man as such. (iii) The sentiment of humanity determines the particular direction which the emotions of approval and disapproval take in human beings. It is because the happiness of men is, as such, pleasing to most men that most men feel approval for qualities which they believe to be pleasant or conducive to human happiness. And it is because the unhappiness of men is, as such, displeasing to most men that most men feel disapproval for qualities which they believe to be unpleasant or conducive to human misery. (iv) The emotion of approval is itself pleasant and that of disapproval is unpleasant.

Of course Hume admits that the sentiment of humanity is often inhibited and overpowered in particular cases. The special relations in which I stand at a certain moment to a certain other man or group of men may completely inhibit

the expression of the sentiment of humanity, which is concerned with them simply as human beings. This obviously happens in the case of jealousy, in war, and so on. Hume also admits that humane emotion may be felt without leading to humane action. All that he asserts is that, in the absence of special causes which excite conflicting sentiments, nearly all men do feel pleased at the thought of a fellow-man in a state of happiness and pained at the thought of a fellow-man in a state of misery. And this seems to be true.

Granted that there is this sentiment of humanity, does it explain the particular direction which the emotions of approval and disapproval take in men? I cannot see that it does. Either the sentiment of humanity is the same as the moral sentiment, or it is not. If it is, then the explanation is merely verbal. This one sentiment is called "the moral sentiment" because it expresses itself in emotions of approval and disapproval, and it is called "the sentiment of humanity" because of the particular direction which these emotions take in men. And, in any case, this identification does not seem to be plausible. To feel moral approval is not the same as to feel sympathetic pleasure, and to feel moral disapproval is not the same as to feel sympathetic pain. Let us then take the other alternative, viz., that there are two different sentiments. If we confine our attention to the positive terms in our pairs of opposites we have now three distinct factors, viz., moral approval, sympathetic pleasure, and something believed to be pleasant or useful to man. The fact to be explained is that the first is directed to the third. The fact alleged as an explanation is that the second exists and is directed to the third. But this explains nothing unless it be assumed that the direction of the first must always be determined by that of the second. And this, whether

true or false, is just as ultimate, and as much or as little in need of explanation, as the original fact which we set out to explain. I cannot help thinking that there is here a latent trace of egoistic psychological hedonism in Hume's theory. I suspect that he is tacitly assuming that the fact that I direct a certain emotion on to the supposed pleasure or pain of *another* is intelligible if and only if it be mediated by a feeling of pleasure or pain in *myself*.

Hume has now to defend his theory on three fronts. (1) Against those who would question his identification of what is generally approved with what is believed to be pleasant or conducive to human happiness. (2) Against egoists, like Hobbes and Spinoza, who would object to his postulating an innate sentiment of Humanity or Benevolence, and would claim to be able to explain all the facts on purely egoistic principles. (3) Against those moralists, whom we may roughly classify as "Rationalists," who would altogether reject his analysis of ethical characteristics, and his view that we can and must determine what kinds of thing are good by ordinary observation and empirical generalisation. We will now consider these three points in turn.

(1) Hume sees that the most plausible objection to his identification of what is generally approved with what is believed to be pleasant or conducive to human happiness arises over legal justice. A particular act of justice may be extremely unpleasant to the agent, who may have to deprive his friend of something which the latter values. It may be extremely unpleasant to the person on whom it is exercised. And it may be detrimental to the general happiness. All these conditions might be realised in carrying out the provisions of a will which was correct in point of law. Yet

we should certainly approve of those concerned if they acted in accordance with the law, and disapprove of them if they did not. Hume's general solution of the difficulty is as follows. If we confined our attention to this particular act and its immediate consequences we should disapprove of it. But, as rational beings, we cannot confine our attention to this very restricted object. We shall inevitably tend to think of its remoter consequences, of the consequences of acts like this becoming prevalent, and so on; and our reaction to this total object may be opposite to that which we should make to the more restricted object which is a part of it.

The application of this general principle to the special case of legal justice is as follows. The happiness of mankind is enormously increased on the whole by there being a set of acknowledged and rigidly enforced rules about the ownership, exchange, and bequest of property. Whatever set of rules be established there will be certain cases in which the enforcement of a rule will lead to worse results than a breach of the rule, if that breach could be taken in isolation. But a breach of an established rule never can in fact be taken in isolation. The whole utility of having rules depends on the fact that they are known to be invariable; and, if you begin to make exceptions in hard cases, this utility will very soon vanish. Any set of rules about property, however arbitrary, so long as it is generally understood and rigidly enforced, ensures greater happiness than no rules at all or rules which cannot be relied upon.

Hume supports this doctrine of the purely utilitarian sanction for legal justice by the following considerations. It is easy, he says, to conceive of circumstances under which rules of property would be useless; and we see, on reflection, that in such circumstances all obligation to keep

the rules would cease. Three such cases can be imagined. (i) Where there is an unlimited supply of goods available to every one, as there is of air under ordinary conditions; or where benevolence was unlimited in extension and intensity. (ii) Where there is such an extreme shortage of goods that, if they were equally divided, no one would have enough to be of any use to him. An example would be a ship-wrecked crew with one biscuit. (iii) Where it is certain that others will disobey the rules, and there is no authority to enforce them. An example would be if one were a member of an army which had got out of hand and was retreating in disorder. The actual position in ordinary life differs from all these extreme cases. There is a limited supply of goods, which is enough for all if properly distributed, and which can be increased or diminished by human action. And men are neither perfectly benevolent nor completely selfish. Under these conditions the existence and enforcement of a set of rules about property is of the utmost utility. A breach of these rules is then in general a double injury. It is always a public injury, as tending to upset confidence in a system whose whole utility depends on the confidence which is felt in it. And in most cases it is a private injury, in so far as it disappoints some man's legitimate expectation of continuing to hold such property as is guaranteed to him by the rules of his society.

Hume argues that the only alternative to his theory is that there is a natural instinct about property. This he denies on the ground of the extreme diversity of the rules about property and the extreme complexity of the notions of ownership, inheritance, contract, etc. No single instinct will account for these facts. But the principle of utility accounts both for the diversities of the rules about property

in different times and places, and for what we find common to all of them. On the one hand, men at different times and places are in very different situations, and so rules about property which are useful in one state of society may be hurtful in another. And, on the other hand, the fundamental needs of men are always the same, and the general conditions imposed by Nature on their fulfilment are fairly constant. This contention, I think, may show that no instinct would be *sufficient* to account for the rules about property, and that real or fancied utility must play an important part. But it does not show that such an instinct may not be *necessary* to account for the facts. The rules about marriage are as odd as, and even more complicated than, those about property; and Hume's argument, if valid about property, ought to show that the rules of marriage have nothing to do with the sexual instinct.

Justice, Hume says, is a virtue natural to man, in the sense that our approval of justice is the inevitable reaction of a being who is both rational enough to consider the remote consequences of acts and benevolent enough to approve of human happiness. And rationality and benevolence are part of the nature of man, in the sense that they are part of his innate constitution. Again, justice is certainly not conventional, if this means that it presupposes an original deliberate contract made among men when they founded societies. For an essential part of justice is the keeping of contracts, and so it is circular to deduce justice from an original contract. It is conventional or artificial only in the sense that there is no need to postulate a special instinct for setting up rules about ownership or a special sentiment which makes us feel disapproval at breaches of such rules. The obvious utility of having rules of some

kind about ownership, and of rigidly enforcing them, fully explains why men have established them and why they feel strong disapproval at breaches of them. But in the details of the rules at any given time and place there is much that is conventional, traditional, and fanciful.

Hume's theory of Justice thus resembles Spinoza's, except that it is not purely egoistic, and that it is more fully worked out. Is it adequate? In the first place, it applies at best only to a small part of justice. It professes to account for our approval of the rigid enforcement of an existing set of rules and for our disapproval of breaches of it. Plainly this is not the whole of the matter. We say that one set of rules is, on the whole, "more just" than another. And we may propose to alter some of the existing rules on the ground that they are "unjust". Now the question whether one set of rules is juster than another seems to be quite different from the question whether the former makes on the whole for greater human happiness than the latter. It seems quite conceivable that one set of rules for distributing property might be far less just than another, and yet that the first might stimulate production so much more than the second that a community would be happier if governed by the first. And I believe that people who were faced with the alternative of introducing one set or the other, or of changing from one to the other, would hesitate between them. For we approve both of justice and of human happiness, and when the two conflict our feelings are mixed.

In this connexion I must add that I question Hume's doctrine that where the utility of justice vanishes our approval of it vanishes too. The truth seems to me to be rather as follows. Where justice and utility conflict, as

they may, our feelings are mixed because we approve of both. And cases may arise in which the sacrifice of justice produces so much human happiness or obviates so much human misery that our total reaction is predominantly one of approval. But, where justice has neither utility nor disutility, as in the case of the ship-wrecked sailors with a single biscuit which is not enough to keep even one of them alive, I think it is plain that we should approve of a just distribution of the biscuit and disapprove of a bestial scramble for it. We should all hope that, if we had to starve along with others, we should have the grace to starve decently and in order, and that they would do likewise.

Again, although I heartily agree with all that Hume says about the extreme utility of having rules of some kind about property and strictly enforcing them even in "hard cases", I am very doubtful whether this fact suffices to explain the original establishment of such rules or the strong feeling of disapproval which we now experience when we contemplate a breach of them. As regards the original establishment of rules about property, it is hard to believe that rather remote and abstract considerations about the happiness of the community as a whole and in the long run would have occurred to the minds of primitive people, or would have had much influence on their conduct unless they had been reinforced by other beliefs and emotions of a less refined kind. As regards our present obedience to such rules in cases where we might profit and escape punishment if we broke them, it seems to me that, if the question of utility comes in at all, it is reinforced by a consideration of justice in the sense which Hume's theory ignores. When I am tempted to do such an act, the question that arises in my mind and sometimes prevents me is this: "Is it

fair that you should enjoy the advantages which you do, through other men keeping the rules when they would profit by breaking them, whilst you take the liberty to break them when it is to your private advantage to do so?"

My conclusion then is that Hume's theory of Justice, though it contains much that is true and important, is inadequate. In particular he has failed to answer the objection that our approvals and disapprovals are in part determined by other considerations beside the supposed immediate pleasantness or unpleasantness, utility or disutility, of the object which we are contemplating. Not only the total amount of happiness to be distributed, but also the way in which it is distributed, stirs our emotions of approval and disapproval. And, although the latter may have a profound influence on the former, that is not the only or the main reason why it arouses the moral sentiment.

(2) We can now pass to Hume's defence of his doctrine against psychological Egoists, like Spinoza and Hobbes. The classical refutation of psychological egoism is contained in the works of Bishop Butler, and Hume does not add much to it. But it will be worth while to give a brief account of his arguments, since later writers of great pretensions, such as Green and Bradley, have been psychological egoists, though not psychological hedonists, in spite of Butler and his refutations.

We may divide Hume's contentions into two groups: (i) Positive evidence in favour of his theory, and (ii) a challenge to his opponents. (i) The positive evidence is as follows. (*a*) It is certain that we feel approval and disapproval of actions and sentiments which we know cannot

affect our happiness at all; *e.g.*, the actions of historical persons in the remote past or of fictitious characters in novels or plays. Again, we may approve of the virtues of enemies, although we know that these very virtues make them more dangerous to ourselves. (Hume lived before the gutter-press had shown us a better way.) Now this must be due either to a direct approval of certain types of character and action, as such; or to a direct approval of human happiness in general, combined with the belief that these types of character and action tend to produce it, even though they affect our own happiness adversely, if at all. Either alternative is inconsistent with psychological egoism. Nor can the facts be explained by saying that we imagine ourselves to be contemporary with the historical characters, or that we imagine the fictitious characters to be real and capable of affecting our happiness. Mere imagination can, no doubt, produce emotion; but it will not continue to do so when we know all the time that it is *mere* imagination, and that the facts are otherwise. (*b*) It is quite certain that we feel approval of qualities which are agreeable or useful to their possessor, even when they are not useful to anyone else. *E.g.*, we approve of a good taste in literature or painting even in a poor man who cannot be a patron of the arts. How can this be explained on egoistic principles?

(ii) The challenge is as follows. On the face of it there is such a sentiment as disinterested benevolence, and the egoist must account for this appearance. He may try to do this in two ways. (*a*) He may suggest that the appearance is due to *deliberate* fraud. This alternative Hume rejects as plainly superficial. We might perhaps add that, if every one knows perfectly well that there is no such thing as

disinterested benevolence, it would not be worth anyone's while to pretend to be benevolent. So we pass to the second alternative, which Hume calls "the more philosophical view". (*b*) This view is that we unwittingly deceive ourselves by some trick of the imagination, some association of ideas, or some bit of mistaken reasoning, when we think that we are feeling an interest in anything but our own happiness. On this theory Hume makes the following comments.

(*a*) Even if it were true, the common distinction between selfish men and actions, on the one hand, and unselfish men and actions, on the other, would correspond to a fact. Granted that in all cases self-interest were the only motive, we must still admit that in some men a certain association of ideas or trick of the imagination or mistaken reasoning causes them to do actions which benefit others rather than themselves. Such men and such actions would be called "unselfish", and it would be a fact that we approve men who habitually deceive themselves in this way, and disapprove those who do not.

(*β*) The affection of animals for each other and for their masters, the love of parents for their children and of men for their friends, are instances of emotions which clearly cannot be reduced to disguised self-interest. There are two comments to be made on this. In the first place, granted that these emotions cannot be reduced to self-love, they are also certainly not instances of general benevolence or humanity, in Hume's sense. They are instances of what Butler calls "particular propensities". They might be admitted to exist, and to be irreducible to self-love, by a man who denied the existence of a sentiment of general benevolence. Secondly, the case of animals and young children would at most prove that apparently disinterested

affection cannot be explained by self-interest and mistaken *reasoning*. It does not prove that self-interest and certain non-rational causes, such as association, might not be adequate to explain the facts.

(γ) He quotes with approval Butler's contention that the possibility of gratifying self-love presupposes the existence of desires for other objects beside one's own happiness. *E.g.*, a revengeful man gratifies his self-love by gaining the pleasures of revenge. But revenge would give him no pleasure if he did not already want to injure his enemy. And this is not a desire for his own happiness, but a desire for another's misery. Hume's argument here appears to be this: "You must admit that we do directly desire some other things beside our own happiness, *e.g.*, the misery of our enemies. If so, why should you deny that we may directly desire the happiness of mankind in general?" This is a valid *argumentum ad hominem* against the psychological egoist. It does not of course prove that we do in fact directly desire the happiness of mankind in general; but it does refute the only argument produced by egoists to show that we do not. For their only argument against the existence of general benevolence is that we cannot directly desire *anything* but our own happiness; and the example of revenge shows that this general principle is false.

(δ) Hume's last argument is characteristically ingenious and plausible, but I believe it to be fallacious. It is this. Not only *has* egoism failed in the past to explain the facts which appear to refute it; we can be confident that it will be no more successful in the future. In physics very familiar phenomena are often found to be due to very complex and previously unsuspected causes. But in psychology "the simplest and most obvious cause that can be assigned for any phenomenon is probably the true one". Strong feelings cannot be accounted for by elaborate trains of reasoning. I may feel very strongly about the death of someone who could not possibly have done me any services if he had lived. Self-sophistication might account for my overlooking the presence of self-interest when it is mixed with other motives, but it cannot manufacture strong feelings out of self-interest where, as in the present case, this motive plainly does not come into operation.

I think that there is a tacit assumption and a confusion in this argument of Hume's. The tacit assumption is that all fundamental emotional and conative dispositions which a man owns must be open to introspection by him simply because they are *his*. If this were true there could of course be no question of a mental occurrence being due to some fundamental tendency which we have never yet recognised. But I see no reason to accept the premise. There might be dozens of fundamental tendencies in ourselves which we cannot detect by introspection, just as there is minute structure in matter which we cannot detect by sense-perception. And what cannot be introspected may cause what can be introspected, just as what cannot be perceived by the senses may cause what can be so perceived.

The ambiguity is this. When it is said that strong feelings can never be accounted for by subtle reasonings, this may mean one of two things. It may mean that a strong feeling in A can never be wholly due to a subtle process of reasoning in A's mind. This is no doubt true. And, in any case, strong feelings which are apparently not egoistic are certainly felt by people who are quite incapable of subtle reasoning, whether valid or invalid. But it might

mean that B's theory about the causation of A's strong feeling cannot be true if it involves subtle reasoning on B's part. Now I see no reason to accept this. It is obviously possible that the causes of A's strong feeling may be very complex and obscure. In that case any correct theory about the causation of A's strong feeling will necessarily involve subtle reasoning on B's part. The upshot of the matter is this. Any egoistic theory which assumes that apparently non-egoistic emotions are *caused by* a subtle process of reasoning in the mind of the *experient* are certainly false. But we cannot reject off-hand an egoistic theory merely because it asserts that apparently non-egoistic emotions are due to very complex *non-rational* causes which need for their detection and analysis very subtle reasoning on the part of the *psychologist*.

My general conclusion on this whole topic is that psychological egoism is certainly false, and that Butler and Hume between them have refuted it and all the arguments which have been alleged in its favour. But to refute psychological egoism is not the same as to prove that there is a sentiment of general benevolence or humanity. I think it very likely that there is such a sentiment; but I doubt whether Hume has proved that there is.

(3) We come now to what is, from the standpoint of ethics, the most fundamental question of all, viz., "Is Hume's analysis of ethical characteristics correct, and is he right in holding that all general rules about what kinds of thing are good or bad can and must be established by observation and empirical generalisation?" Hume discusses this question in the form: "What are the respective functions of Reason and Feeling in ethical matters?"

Unfortunately he never explicitly says what he means by "Reason". Now "Reason" is a highly ambiguous word, and I suspect that Hume uses it in this discussion in an unduly narrow sense. It will make for clearness if I state what I understand by "Reason" before I begin to deal with Hume's arguments. I ascribe three cognitive functions to Reason: (i) The intuiting of necessary and universal connexions between characteristics, when conjunctions of these characteristics are presented to the mind's attention. *E.g.*, it is an act of Reason, in this sense, when we see by inspection that *any* triangle which is equilateral *must* be equiangular, and conversely. In this way we derive our knowledge of axioms. (ii) The drawing of inferences, demonstrative or problematical, from premises. This activity is, no doubt, closely connected with the former. For it depends on seeing certain formal relations between propositions, and on recognising that such relations justify inference in *any* instance in which they are present. (iii) The formation of *a priori* concepts. This needs explanation. It appears to me that we have concepts of certain characteristics which are neither manifested to us in sensation (as *redness* is) nor synthesised out of characteristics so manifested (as the characteristic of *phœnixhood* is). I believe the concept of *Cause*, and many others, to be of this nature. I have no doubt that certain specific kinds of sensible experience are necessary conditions for the formation of such concepts; but they are not, strictly speaking, derived from sensible experience, as the concepts of *redness* and *phœnixhood* are. These are what I call "*a priori* concepts". Some people would deny that there are any such concepts; and those who would admit them might differ very much about their nature and status. If there be

a priori concepts, as I believe there are, I ascribe the formation of them to Reason. The three cognitive functions which I assign to Reason may be called respectively "Intuitive Induction", "Ratiocination", and "Formation of *A Priori* Concepts". Now it is an essential principle or prejudice with Hume to deny the possibility of *a priori* concepts; so naturally he does not include the third function under the head of Reason. But in his other works Hume does admit Intuitive Induction; for this is involved in what he calls "knowledge of the relations of ideas" and contrasts with "knowledge of matters of fact". Yet here, it seems to me, he ignores this function of Reason altogether, and tacitly reduces Reason to Ratiocination. We are now in a position to consider his arguments.

Hume's general position is the following. The *prima facie* case for the man who thinks that Reason plays an essential part in ethical matters is that we certainly do dispute about questions of right and wrong, and do try to persuade each other on moral questions. Now we do not dispute about mere feelings and emotions. The *prima facie* case for the man who thinks that sentiment and emotion play an essential part in ethical matters is that virtue and vice certainly do move our feelings, and that moral approval and disapproval are undoubtedly motives to action. Now Reason cannot tell us that one quality must attract and another must repel us. This must depend on innate or acquired tastes. And the mere intellectual recognition of the presence or absence of a certain quality or relation neither moves our feelings nor affects our actions.

He concludes that Reason and Sentiment both play an essential part, but that the parts are quite different. Reason is needed to tell us that certain types of character or conduct

tend to produce happiness or misery in the agent or in other men. When the situation is complex and the consequences are mixed, Reason is needed to analyse the situation and to estimate the balance of happiness or misery which is likely to result. But this knowledge which Reason gives us would lead neither to approval nor disapproval, action or abstention, unless the thought of human happiness attracted us and the thought of human misery repelled us. Now this attraction and repulsion cannot be due to Reason, but must depend on the special emotional make-up of the human mind. The essence of Hume's view then is that Reason is wholly confined to matters of fact. It will help us to analyse a situation, to choose means for a given end, and to infer probable consequences of various alternative courses of action. But it has nothing whatever to do with our choice of ends as distinct from means. We desire things as ends only because they move some emotion in us, and not because of any objective characteristic in them which Reason can recognise.

It is evident that there are two different propositions involved in Hume's doctrine. The first is that Reason, even if sometimes necessary, is never sufficient to account for the facts of moral emotion and moral action; and that a Sentiment must be postulated in addition to explain these. The second is that Reason is concerned only with matters of fact. Now the first of these contentions may be, and, I believe, is true. But it is little more than a truism; and it has no tendency to support the second proposition. Suppose it were the case that there is a certain quality, viz., goodness or badness, and certain relations of rightness or fittingness, which are recognised by Reason and by it alone. It is still logically possible that a being who was

rational in the cognitive sense, *i.e.*, who recognised these qualities and relations, should be entirely unmoved by the thought of their presence or absence. And it is logically possible that a being who recognised these qualities and relations and felt emotions of approval and disapproval when he thought of their presence or absence should not be moved to do what he approves or to avoid what he disapproves. No doubt we should call such beings " moral lunatics ", and say that they are " not completely rational ". But the fact that they are conceivable, and that they do indeed exist, shows that even the most convinced Rationalist about moral *cognition* must postulate certain emotional and conative dispositions in addition to Reason in order to account for moral *feeling* and moral *action*. Now some Rationalists have written as if they thought that the mere recognition of ethical characteristics by Reason *sufficed* to account for moral feeling and moral action. If any of them really did think this, they were wrong ; and Hume's argument shows that they were. But this has not the faintest tendency to prove that they were wrong in holding that Reason is *necessary* for the recognition of ethical characteristics and for the intuiting of necessary connexions between them and other characteristics. Thus the second part of Hume's contention, viz., that the only business of Reason is with matters of fact, is quite unsupported by the excellent reasons which he gave for the first. Is there any reason to believe it ?

Hume never states very clearly the alternatives to his own theory. I think it will be wise to do this before considering in detail his arguments for it and against its rivals. Let us grant, for the sake of argument, that the judgment " X is good " would never have been made *in the first*

instance unless the person who made it had felt an emotion of approval in contemplating X, though it may *now* on occasion be made by a person who is not feeling this emotion. This may be compared with the fact that the judgment " X is red " would never have been made in the first instance unless the person who made it had had a sensation of red on looking at X, though it may now on occasion be made by a person who is not having such a sensation. Now there are two different ways of analysing the latter fact. The first would be to say that " X is red " means simply " Most men will have a sensation of red when they look at X." This may be called the " phenomenalist analysis ". The second would be to say that " X is red " means " There is a certain property in X which causes sensations of red in most men who look at X." This may be called the " causal analysis ". Let us now apply this to the case of goodness. The phenomenalist analysis would be that goodness is the characteristic of *being* generally approved by men. The causal analysis would be that goodness is the property which *causes* a thing to be generally approved by men. It is plain that Hume takes the phenomenalist view about goodness. According to him the property which causes a thing to be generally approved by men is not goodness but supposed direct pleasantness or utility. But he has produced no conclusive reason for preferring the phenomenalist to the causal analysis.

We have now to consider another alternative. As before we will begin with a parallel from non-ethical topics. It is generally held that the judgment " X causes Y " would not have been made in the first instance unless a number of X-like events had been observed and they had all been found to be followed by Y-like events. The phenomenalist

analysis of this fact is that " X causes Y " simply means " X-like events will always be followed by Y-like events." But another view is possible. It may be that there is a peculiar relation between X and Y which cannot be manifested through the senses, but which is intuited by the intellect when and only when a number of sequences of X-like and Y-like events have been presented to its attention through the senses. This of course makes the causal relation an *a priori* concept, in the sense defined above. I will therefore call this type of analysis the " *a priori* concept analysis ". Now the *a priori* concept analysis of the ethical fact which we are granting to Hume would be as follows. Emotions of approval and disapproval furnish the necessary occasions on which the intellect intuites certain ethical relations, *e.g.*, those of rightness and wrongness, fittingness and unfittingness, which cannot be manifested through the senses. We could not expect Hume to entertain this suggestion, but it is nevertheless a perfectly possible one.

We are now in a position to consider Hume's arguments. He has two arguments against the Rationalist's position, and three in support of his own. (i) Rationalists maintain that actions, intentions, or emotions are right or wrong because of some relation of fittingness or unfittingness to something else, which Reason recognises. Hume says that this relation must either relate the action or emotion to the situation in which it takes place, or it must be the logical relation of falling under or conflicting with some general moral rule. If the former is meant, he challenges the Rationalist to point out exactly what this relation is. If the latter is meant, he argues that the theory is circular. For the general moral rule must have been reached by induction from observed particular cases of right and wrong

actions. Particular actions must therefore be recognised to be right or wrong before any general moral rules could have been formulated. Hume's challenge seems to me unfair. Might not the relation in question be absolutely unique and peculiar, and yet perfectly familiar? If so, any attempt to express it in other terms would necessarily be erroneous or tautologous. On the other hand, his objection to the second form of ethical Rationalism seems fairly conclusive.

(ii) Inanimate objects may have to each other exactly the same kind of relations which would make us approve or condemn human beings. Yet we do not make ethical judgments about inanimate objects. When a young tree destroys the older tree which produced it, the two trees stand in precisely the same relations in which Nero and his mother stood when he murdered her after she had gained him the empire. Yet we blame Nero, and do not blame the young tree, for ingratitude. I do not think that a Rationalist need spend many sleepless nights over this objection. Nero and his mother had minds, whilst we believe that the trees had not. In virtue of this difference Nero and his mother stood in mental relations in which the trees could not have stood. And we condemn Nero in respect of his emotions and intentions towards a person who had had certain emotions and intentions towards him.

We come now to Hume's three arguments for his own view. (i) In geometrical reasoning we first observe certain relations between points, lines, etc., and then proceed to *deduce* other relations which were not before obvious to us. But, when we reflect on a situation in order to pass a moral judgment, *all* the relations must be known *before* we can pass the judgment. Thus Reason must have completed its task before moral judgment can begin, and its task is simply to

ascertain the exact facts of the case. All that then remains is for the situation which Reason has analysed to call forth an emotion of approval or disapproval. There are two undoubted truths in this argument of Hume's. (*a*) I must be fully aware of the non-ethical relations in a situation before I can make a trustworthy judgment on the ethical relations. (*b*) When I am fully cognisant of the non-ethical relations I cannot infer, from them and them alone, the ethical relations; as I might seem to infer the remaining geometrical relations between a set of points from a selection of their geometrical relations. But, even in the geometrical case, I do not infer the additional geometrical relations *simply* from those which are already known. I infer them from these *together with* the axioms of geometry, which are known by Intuitive Induction. Similarly, it is arguable that I first recognise the co-existence of certain non-ethical relations with certain ethical relations in a particular case; then see by Intuitive Induction that the presence of the former *entails* that of the latter in *any* case; and finally use this as a premise for inferring the presence of these ethical relations in other cases in which I find these non-ethical relations. So the premises of this argument are quite compatible with the view that Reason plays a much more important part in ethics than Hume allows.

(ii) Hume argues that his position is strengthened by the analogy between ethical and æsthetic judgments. The beauty of an object no doubt depends on the relations and proportions of its parts. And these are in many cases recognised only by the exercise of Reason. But the recognition of these relations and proportions is not sufficient to give rise to an æsthetic judgment. A circle would have no beauty unless there were observers so constituted that

the recognition of its form calls forth an emotion of admiration in them. Similarly a murder would not be wrong unless there were observers so constituted that this kind of relation between men calls forth an emotion of disapproval. This argument does not, I think, appreciably strengthen Hume's position. Either the situation with regard to æsthetic judgments is, or it is not, exactly analogous to that with regard to moral judgments. If there is exact analogy, we have already shown that the facts in the case of moral judgments are susceptible of two other interpretations beside Hume's. And the same two alternatives would be open in the case of æsthetic judgments. If there is not exact analogy, then the argument from the æsthetic to the moral judgment cannot be relied upon. For the differences might be such as to allow Hume's theory to be true of æsthetic judgments, and to prevent it being true of moral judgments.

(iii) If you press a man as to why he did a certain action there will always come a point at which he can make only a tautologous answer. If you ask him why he plays golf, he may say that it is for the sake of health. If you ask him why he wants to keep in health, he may say that it is because illness is painful. But, if you ask him why he dislikes pain, and he still has patience to answer you at all, he can only make the tautologous answer: "Because I do." This, Hume thinks, shows that Reason is concerned only with means and with relative ends, never with ultimate ends. Now virtue is admitted to be an ultimate end, desirable for its own sake. Therefore there must be some sentiment in men to which virtue appeals, and it must derive its value from this and this alone. The weakness of this argument will best be seen by taking a parallel case. In any chain of reasoning whatever we eventually get back

to premises for which we can give no reason, in the sense that we cannot mention any other proposition from which they are deducible. But this does not show that our acceptance of these ultimate premises must be irrational. It may of course happen to be so. But it may be that we accept them because Reason perceives directly that their subjects and their predicates are necessarily connected. Similarly, in explaining why we acted in a certain way, we come eventually to ends which are valued for themselves and not as means to anything else. But it does not follow that our recognition of their value does not depend on rational insight into their nature.

The upshot of the matter is that, on this vitally important point, Hume has neither proved his own case nor refuted that of his opponents. But it remains possible that he is right and they are wrong. I cannot profess to decide the question here; but I will end by pointing out one consequence of Hume's view. This is that every dispute on questions of right and wrong is capable of being settled completely by the simple method of collecting statistics. Suppose that A thinks that X is right, and B thinks that X is wrong. We have first to make sure that A and B agree as to the non-ethical facts about X, *i.e.*, as to its non-ethical qualities and relations to other things, as to what effects it will have and what effects other things which might have been substituted for it would have had, and so on. Suppose that, when all differences and confusions on these non-ethical matters have been removed, A still thinks that X is right and B still thinks that it is wrong. If Hume's theory be true, this means that A thinks that most men would feel an emotion of approval on contemplating X, whilst B thinks that most men would feel an emotion of disapproval

on contemplating X. Now this is a question which can be settled by experiment, observation, collection of statistics, and empirical generalisation. This seems to me simply incredible. I should accept the view that there is a point in any ethical dispute between A and B beyond which further argument becomes futile. This would not, of course, prove that the difference has been reduced to a mere difference of taste; for it might be that A's intellect was obtuse or warped, as compared with B's, in respect to certain quite objective qualities or relations. But, as I have just pointed out, the logical consequence of Hume's theory is not that in disputes on moral questions there comes a point beyond which we can only say "*de gustibus non est disputandum.*" The logical consequence of his theory is that all such disputes *could* be settled, and that the way to settle them is to collect statistics of how people in fact do feel. And to me this kind of answer seems utterly irrelevant to this kind of question. If I am right in this, Hume's theory must be false.