

DAWES HICKS LECTURE

THE NATURALISM OF BOOK I OF HUME'S
TREATISE OF HUMAN NATURE

BY DAVID PEARS

Read 14 October 1976

HUME does three things in the account of the understanding that he gives in the *Treatise*. First, he argues that certain concepts and beliefs cannot be explained or justified as the products of reason operating on sensory experience, and that they are always misrepresented by philosophers who treat them in this way. His second step is to consult the vulgar, or ordinary people, in order to find out how they regard the concepts and beliefs, and never afterwards to stray too far from their view of them. Finally, he offers his own theory, which explains and sometimes justifies them. The first of these three stages is destructive, and it clears the ground for the construction of an ambitiously conceived alternative both to rationalism and to naïve empiricism in the later stages. The alternative is a blend of sophisticated empiricism and naturalism. The contents of the mind are all traced back to original experience, which he describes without claiming to be able to explain it, and then he offers a naturalistic explanation of what the mind does with this material.

Anyone who works his way through this system wondering whether it leads to some form of scepticism will be surprised to find that the outcome is uncertain. This is partly because the second stage of Hume's account of the understanding is not purely descriptive, but contains some reinterpretation of the concepts and beliefs of the vulgar, trimming them to fit the theory that will be offered in explanation of them in the third stage. In fact, his attitude towards ordinary people is ambivalent: he treats them respectfully when he is criticizing rationalism, but, when he is developing his own naturalistic theory, he sometimes finds that they indulge in excesses that cannot be accommodated within it and, therefore, have to be trimmed and reduced. Many philosophers hold that his treatment of causal necessity illustrates this tendency. But if such reductions are legitimate, what are the criteria of scepticism?

Should it be judged by a failure to accommodate the whole system of ordinary people, or only its irreducible core?

Another reason for the uncertainty of the outcome is doubt about what counts as a justification. According to Hume, a concept is justified when it is derived from available experience. But the restrictiveness of his rules of derivation is evidently questionable. However, his own misgivings are mostly concerned with the justification of beliefs, and they arise in two different ways. First, if a belief is the conclusion of an essentially fallible inference, he is inclined to doubt whether the psychological inevitability of that type of inference is a sufficient justification of it, and that is one route to scepticism. Second, when he develops his own naturalistic explanation of fallible inferences, he finds it hard to discriminate between the sensible ones and the foolish ones. This is a fault in his theory, because it ought to have split the class of fallible inferences in a way that at least corresponded to the division into the sensible and the foolish.

He is, of course, aware which of the inferences allowed by his theory are sensible and which are foolish. For when he applies it to ordinary people's beliefs about the objects of perception, he knows that he is giving them the benefit of its lack of discrimination. He interprets the inferences that lead to their beliefs in a way that makes them out to be foolish and flighty, when they are viewed from his Cartesian standpoint, but he explains their origin by pointing out that his theory, or at least his application of it, allows foolish flights of fancy of precisely that kind. I shall not concern myself with his interpretation of the ordinary, pre-theoretical belief that objects continue to exist unperceived, though he certainly exaggerates its blindness to theoretical issues. Nor shall I examine his Cartesian assumption that the causes of our sensory impressions can only be inferred. The point for which I shall argue is that his theory ought to have provided some basis for the distinction between sensible and foolish fallible inferences, and that, if it had done so, it would have been clearer whether it justified any of them, and so whether it avoided scepticism.

I am not going to pursue the question whether Hume's naturalistic system is sceptical in outcome. I am using it only as an introduction to my subject, which is the naturalism itself. But I shall not disguise my conviction that an adequate account of the origin of a belief can no more fail to contribute to its justification than an adequate account of the origin of a concept can fail to make a similar contribution.

The naturalism of Hume's account of the understanding goes deeper in the *Treatise* than in his later writings. In the *Treatise* he not only describes our habits of thought but also tries to deduce them from a single theory in the manner of Newton. The theory deals with what he calls 'natural relations between ideas',¹ and most of it is based on psychological observation. But the theory also contains two axioms which are not based on psychological observation, though he sometimes writes as if that were their basis: '*All our ideas are derived from impressions*',² and '*All our distinct perceptions are distinct existences*'.³ In his later work his naturalism lacks this depth, because he gradually abandons his elaborate psychology and describes our habits of thought without trying to fit them into a single explanatory framework. Apparently, he was dissatisfied with the foundations laid in *Book I*, but did not see how to improve them and wanted to complete the edifice.

It does not follow that his later account of the understanding is shallow. Certainly, it lacks theoretical depth, but there are other ways in which naturalism can achieve profundity, and Wittgenstein may have been right in thinking that theoretical depth is for scientists rather than philosophers. Anyway, Hume's later naturalism is more informal, but it is the result of penetrating observation of our habits of thought, and it subjects them to a profound critique based on the two axioms, which, of course, he never abandoned. But his early naturalism is systematic, and that is what gives it its depth and interest.

The early psychological theory is divided into three parts. First, there is a very economical account of the contents of the mind: it contains only impressions and ideas that are derived from them and copy them. The second part is the deceptively simple thesis that belief is a vivacious idea. Then the origin of belief is described in the third and most complicated part of the theory, which analyses natural relations between ideas.

I shall concentrate on the second and third parts of this theory, because I am going to inquire whether it provides an adequate explanation of our beliefs about causation, perceived objects, and persons. But it must not be forgotten that the first part contains Hume's theory of meaning. That theory is expressed in an axiom that is not really based on psychological observation. For though the precise ways in which ideas can be derived from impressions may be discernible by personal

¹ *Treatise* I. I. v.

² *Ibid.* I. I. i.

³ *Ibid.* I. I. vii, and Appendix.

reflection on their history, the thesis that in one way or another all ideas must be derived from impressions has to achieve some kind of fit with the distinction between genuine and spurious ideas that was already established and current. Nor must it be forgotten that his system relies at many points on the other axiom, that all our distinct perceptions are distinct existences. This axiom too needs to be supported by something more than direct psychological observation. For it is not the case that Hume tried but failed to discover real connections between distinct perceptions. On the contrary, he began by choosing discrete elements for his theory, and that is the kind of choice that would be vindicated not by his failure to discover elements of any other type, but rather by some kind of global fit between the resulting theory and human thought. These two axioms are permanent features of his philosophy, and his naturalism would have been lame without them.

At first sight, the support received by the thesis that belief is a vivacious idea appears to be purely local. Hume simply tries to report the result of direct observation of the psychological phenomenon in his own case. But he does not find it easy to formulate the result in a satisfactory way, and he vacillates between two versions of it. Sometimes he speaks of *lively colours* and *brightness*, but at other times he prefers *force*, *firmness*, and *steadiness*. In the end he confesses that 'it is impossible to explain perfectly this feeling or manner of conception',¹ and he characterizes it as 'that je-ne-sçai-quoi, of which it is impossible to give any definition or description, but which everyone sufficiently understands'.²

It is fairly clear what his difficulty was. His ideas are images, and, if the belief that a certain kind of impression is about to occur consists in an idea, the idea will need two entirely different properties. In order to carry the content of the proposition that is believed, it must possess an intrinsic property that matches the identificatory property of the expected impression. This requirement is most easily understood in a visual case, because the intrinsic property of the idea may then be a pictorial property of an image. The other property needed by the idea is an extrinsic one which will mark off belief from other propositional attitudes. Now Hume is almost aware that in a visual case the property that marks off belief must not be pictorial, because, if it were pictorial, the propositional content would change

¹ *Treatise* I. 3. vii.

² *Ibid.* I. 3. viii.

as mere consideration turned into suspicion and finally into belief.¹ So it is fairer to discount his tendency to treat the mark of belief as a pictorial property of the idea, and to take his settled view to be that it is a non-pictorial property, or, more generally, an extrinsic property.

However, that is not the end of his difficulty. For when he has decided that the mark of belief must be an extrinsic property of the idea, he still has to identify the right extrinsic property, and that proves to be a difficult task because all the likely candidates seem to occur not only in cases of belief, but also in cases in which belief is not, and sometimes cannot be, involved. For example, it is obvious that an obsessive fantasy may force itself upon the mind and may haunt it steadily and even with firmness. The reason why all Hume's suggestions founder on such counter-examples is that it is impossible to develop a theory of belief that makes so little use of assertion and truth. Force and steadiness are marks of belief only in the context of a question. An image that forces itself upon the mind and maintains itself there as the pictorial answer to a definite question does amount to a belief. But the theory collapses when this context is omitted, and an analysis of the context would have to be based on the concepts of assertion and truth. So it was an illusion to suppose that a theory of belief could be supported by the kind of local evidence that Hume produces.

The theory of natural relations is intended to explain the origin of our problematical beliefs. A natural relation between two mental elements (i.e. impressions or ideas) is one that usually but not invariably sets up an association between them.² A problematical belief is an idea that acquires its vivacity through an association with a present impression.³ Memory is not a problematical case, because its beliefs are ideas whose vivacity is not acquired by any detectable mental process, but by some neural process with which Hume is not concerned.⁴ Perception is not problematical so long as its judgements are confined to present impressions, because the conviction produced by them is too strong to count as belief.⁵ But beliefs about objects that continue to exist unperceived are problematical. So too are beliefs about persons and causal processes.

In Hume's theory there are two different ways in which relations may hold between mental elements. They may relate

¹ *Ibid.* I. 3. vii.

³ *Ibid.* I. 3. viii.

⁵ *Ibid.* I. 4. ii.

² *Ibid.* I. 3. iv-vi.

⁴ *Ibid.* I. 3. v.

them through their contents, or independently of their contents. If two ideas are related through their contents, the same relation may, and sometimes will, hold between any similar pair of impressions. He regards it as a speculative question whether the relation also holds between any pair of objects that continue to exist unperceived. But the match between the related pair of ideas and similar pairs of related impressions provides him with a workable, but rudimentary theory of truth.

Every relation has an intrinsic nature, and, when we consider its intrinsic nature, Hume says that we are taking it as a philosophical relation, because philosophers extend the class of relations beyond the limit that strikes ordinary people as natural.¹ But some philosophical relations are also natural relations, because they lie within the limit that is recognized by ordinary people, and so they tend to produce associations between the mental elements that they relate.² He places only three relations in this sub-class: resemblance, contiguity, and causation.³ The psychological theory of the *Treatise* is deduced from the proposition that these three relations are distinguished from other philosophical relations by the fact that they are also natural relations, and so set up the associations that produce our problematical beliefs.

The first step towards understanding the theory is to see why certain ordinary beliefs are problematical and need it to explain them. This question may be approached through Hume's classification of philosophical relations. When he considers relations strictly from the point of view of their intrinsic natures, he divides them into two classes. In the first class the relation holds necessarily between two ideas given their contents, and yields knowledge rather than belief. Resemblance and certain mathematical and logical relations belong to this class.⁴ In such cases he allows the rationalist inference from ideas to things, even if the things are only interpreted phenomenally as impressions, and this gives him a theory of necessary truth. In the second class the relation still holds between the two ideas through their contents, but it only holds contingently, so that it is an open question whether it also holds between impressions. How can such contingent truths be established?

Hume does not give the same answer for all the relations in this second class, but subdivides it for this purpose. One subdivision contains contiguity while the other contains causality.

¹ *Treatise* I. 1. v.

² *Ibid.* I. 1. iv.

³ *Ibid.* I. 1. iv and 3. v.

⁴ *Ibid.* I. 3. i.

Contiguity may be established by direct observation,¹ but causation cannot be established in this way.² Causal beliefs are, therefore, problematical. He could have given the same explanation of the problematical character of beliefs about persons and unperceived objects. He need not have treated identity as a relation, because he could have used two substitutes for it which would have made his theory much clearer—the co-personality of mental elements, and the co-objectuality of impressions of sensation. Like causation, these two relations evidently do not hold necessarily between mental elements, nor can their holding be established by any special impression.

It may seem surprising that Hume should put so much emphasis on the fact that resemblance, contiguity, and causation are natural relations as well as philosophical relations. For in non-problematical cases it hardly seems necessary to mention this duality of aspect. The resemblance between two ideas identified through their contents is established *a priori*, and though it may set up an association between them, the association evidently cannot play any part in producing the knowledge that they resemble one another, because it is the later effect of that knowledge. Similarly, contiguity is established by direct observation, and, if it sets up an association, it will be one that is subsequent to the formation of the conviction that the two mental elements are contiguous, and cannot, therefore produce it.

But in problematical cases, like causation, the duality of aspect of the three relations is essential to Hume's psychological theory. For when the belief that a relation holds between two impressions cannot be established either as a necessary truth or as a contingent truth based on a third, special impression, he has to explain, not only which features of the original impressions lead us to believe that the relation does hold between them, but also how they produce this belief. The first of these two explanations is provided by the analysis of the relation in its philosophical aspect, which will give the total legitimate content of the idea, or, to express this in another way, all that there can be in reality. The second explanation is provided by the analysis of the relation in its natural aspect. For this will show how the features of the original impressions produce the associations required for the belief.

So Hume's psychological theory is designed to explain the origin of problematical beliefs. But it also contains an account of

¹ Ibid. I. 2. iii.

² Ibid. I. 3. xiv.

the way in which such beliefs achieve the match with impressions that counts as truth in his system. They do not achieve it *a priori*, as rationalists suppose, nor by a direct adjustment to a special experience, as naïve empiricists suppose, but in a third, more complicated way, which he explains most successfully in the case of causation.

He defines causation both as a philosophical and as a natural relation. From the philosophical point of view a cause is 'an object precedent and contiguous to another, and where all the objects resembling the former are placed in like relations of priority and contiguity to those objects that resemble the latter'.¹ If he had been more cautious, he would have set up this definition on a phenomenal basis, substituting impressions for objects. Either way, the total legitimate content of the idea of causation is constant conjunction between resembling pairs of things, in each of which one member is contiguous with and prior to the other.

From the natural point of view, a cause is 'an object precedent and contiguous to another, and so united with it that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other'.² He is sometimes criticized on the ground that these two definitions are not necessarily equivalent, because the second, unlike the first, requires the presence and operation of a mind. But it is clear that the consecutive clause after the words 'so united with it' merely specifies the type of union through its tendency to generate in any mind that may observe it an association of two mental elements, which, on some later occasion will produce the causal belief that something resembling the second element is about to occur.

I shall not discuss the question whether this analysis of causal belief trims it too ruthlessly to fit an excessively restrictive theory of meaning. If it is not too reductive, Hume's account of the internal impression of necessitated transition certainly gives an ingenious explanation of the illegitimate element in the popular concept of causation. But my concern is with the naturalistic theory of belief, which is designed to explain how causation, when it relates a sequence of pairs of impressions, produces a reflection of itself in the mind of the observer. The reflection is a disposition to produce a vivacious idea of the second member of the pair in response to an impression re-

¹ *Treatise* I. 3. xiv. ² *Ibid.*, loc. cit.

sembling the first one. It is, therefore, a causal disposition, and it is caused by the original observed causation functioning as a natural relation.

Hume's analysis of causation, considered as a philosophical relation, is plausible, and his account of the way in which it functions as a natural relation producing a mental reflection of itself is ingenious, and the whole treatment offers a promising way of avoiding the choice between rationalism and naïve empiricism. But the explanation of the origin of causal belief was designed to fit into a general theory, which would also serve to explain other problematical beliefs, and Hume is less successful in carrying out this more ambitious project.

The general theory is that belief is an idea endowed with vivacity through its association with a present impression, and that such associations are produced by the three natural relations, resemblance, contiguity, and causation. When this theory is applied to the analysis of causation itself, it seems at first sight to fit. For observed causation, functioning as a natural relation, produces an association which, in its turn, produces a vivacious idea in response to a present impression. But when observed causation is analysed as a philosophical relation, it turns out that its most important ingredient is constant conjunction. It is true that contiguity is another ingredient, but it is a less important natural relation between the members of each pair in the sequence, and resemblance between the members of each pair is not needed at all. Of course, the first members in each pair must resemble each other, and so too must the second members. But that is a different requirement, because, though this resemblance, like any other, may be regarded as a natural relation, it certainly does not function as such in this case. For it does not set up the association, but only determines the two sets of members which are going to be associated by some other means, viz. by constant conjunction and contiguity.

So there is some confusion in Hume's general theory about the origins of belief. The theory is that resemblance, contiguity, and causation are three equipollent natural relations, each capable of setting up an association that will produce a belief. But in the case of causation it is constant conjunction that does most of the work of setting up the association, and causation is the only one of the three natural relations that involves constant conjunction. Contiguity has a minor role, but it does at least relate the members of each pair in the sequence, while

resemblance does not even do that. It is true that all three natural relations are involved in causal inference, but their contributions are neither equal nor of the same kind. They could not be equal, because causation includes the other two and adds to them something much more weighty as evidence—constant conjunction, and resemblance does not even make a contribution of the right kind. For when a natural relation sets up an association of ideas that produces a belief, it is essential that it should hold between the two things that the ideas match. But in the case of causal inference resemblance does not meet this requirement. Later I shall analyse another, more important case of failure to meet this requirement—a failure that occurs in Hume's account of the ordinary belief that objects continue to exist unperceived.

The confusion in his general theory of the origins of belief is more evident when resemblance and contiguity are taken on their own, apart from causation. For then the weakness of these two natural relations is immediately exposed. The observed contiguity of two things is not enough to support the belief that the next time one of them is found the other will still be found beside it. A general proposition is evidently needed, and it is unlikely to be the proposition that nothing moves. Whatever it is, it will need to be based on some constant conjunction. So contiguity alone is a completely powerless natural relation.

Resemblance is in a worse plight, because it is not even clear what the belief would be in a case in which resemblance was operating alone. If a present impression produces an idea associated with it by resemblance, the idea needs a reference in order to function as a belief. But Hume's theory of belief omits all propositional details, and in this kind of case it is not clear what material could be used to supplement the theory and give the belief a definite reference. On some occasions the belief would be that the next impression of sensation will resemble the present one, and in his account of the ordinary belief in unperceived objects Hume attaches great weight to this kind of constancy. But this belief too needs a general proposition, more subtle than the proposition that nothing changes, and based on some constant conjunction. Incidentally, this kind of inference ought to have led him to add something to his thesis that resemblance holds necessarily between two mental elements given their contents. He should have added that, when a mental element is specified not by its content but

by its place in a sequence, it can only be a contingent fact that it resembles a present impression.

In one way it is unfair to make these points against Hume. For he is aware that the only sensible kind of inference to a new factual conclusion is causal inference, which depends on custom generated by observed constant conjunction. So he admits that resemblance and contiguity, operating on their own, do not produce sensible beliefs, and he actually considers the objection that, if his theory were correct, they would do so.¹ But as Kemp Smith showed long ago,² his answer to the objection is a wholly inadequate defence of his general theory. He merely argues that resemblance between cause and effect sometimes reinforces a causal inference, and points out that a visit to the scene of some supposed historical event sometimes reinforces the belief that it actually occurred. Such observations clearly will not save the theory, and he shows a tendency to modify it in this section of the *Treatise* by saying that a present impression can give an associated idea a liveliness that does not amount to belief. But that would reopen every question.

It is not enough that he should be aware of the fact that the only sensible kind of inference to a new factual conclusion is causal inference. In the *Treatise* he is not content with recording this fact. He is trying to achieve a more systematic kind of naturalism, which will not only list the kinds of inductive inference that we find acceptable, but also relate them to a general theory. Now the general theory might not justify our grading of the evidence, but it ought to explain it, and the explanation should make some contribution towards justifying it. But Hume's general theory does not begin to do these things, because it does not provide any basis for the distinction between sensible and foolish inferences to new factual conclusions.

It is possible to see how the theory lost contact with the phenomena. Association is first introduced to account for the occurrence of ideas, and associations with this function really are produced by each of the three natural relations, resemblance, contiguity, and causation. But when association is given the altogether different function of generating belief, it can be produced only by causation or by some other relation reinforced by constant conjunction. Hume himself points out this fact, but he never revises his general theory to fit it. Perhaps he fails to see the need for revision in the *Treatise*, and perhaps the reason for this is that so much of Book I is taken up

¹ *Treatise* I. 3. ix. ² 'The Philosophy of David Hume', pp. 378-82.

with the application of the theory to causation, which is more successful because causation is the favoured exception.

Since Hume gradually abandoned his general psychological theory in his later works, it might seem appropriate to regard it as an excrescence on the *Treatise*, with no effect on his analysis of problematical beliefs. But this would be a mistake. For, as I shall now show, one of the flaws in the theory vitiates his account of the ordinary belief that objects continue to exist unperceived.

He often expresses this belief as the identity-statement, that what is perceived now is the same as what was perceived earlier. But he is not especially concerned with the criteria for reidentifying perceived objects, any more than with the criteria for reidentifying persons. His main concern is always with the principles of union, which in this case have to span an interruption of perception. Consequently, what he investigates is really a relation, co-objectuality. Now according to him, ordinary people do not distinguish between perceptions (i.e. in this case impressions of sensation) and objects.¹ So when he consults them in order to discover the category of the terms that they take to be related by co-objectuality, the result is unclear. He represents their answer as 'Perceptions'. But this answer is supposed to be given at a pre-theoretical stage, with no distinction drawn between perceptions and objects, and so the inference, 'Therefore, impressions of sensation', may be blocked by opacity. The situation would have been easier to grasp if he had represented their answer as 'What is immediately perceived', or, more briefly, 'Data'.

There is another point on which the beliefs of ordinary people need interpretation. Do they believe that data continue to exist when perception is interrupted, and that, when they are perceived, they would still have existed even if they had not been perceived? Hume credits them with affirmative answers to both questions.² So he does not suppose that they regard unperceived data as mere possibilities in the spirit of phenomenalism. (If they had taken that view, it would not have actually been phenomenalism, because they do not unequivocally identify data with impressions of sensation.) On this point he is on their side. For though he rejects the suggestion that data are specifically different from impressions of sensation,³ he never considers the suggestion that, when they are not perceived, they are mere possibilities.

¹ *Treatise* I. 4. ii.

² *Ibid.*, loc. cit.

³ *Ibid.* I. 2. vi.

If ordinary people believe that the fact, that two detached sequences of perceived data are co-objectual entails that they actually continue to exist in the interval during which perception is interrupted, their belief is obviously problematical. It cannot be a deliverance of reason, because it can be denied without any change in the contents of the ideas involved in it. It cannot be based on sensation unaided by inference, because no impressions of sensation can possibly support hypotheses about something that is not perceived.¹ If neither rationalism nor naïve empiricism can account for it, the explanation must be sought in Hume's general psychological theory of the origins of belief. Co-objectuality must be analysed as a philosophical relation, and its analysis must include natural relations capable of producing associations to sustain the ordinary belief in it.

Before examining Hume's attempt to carry out this task, it is as well to state the conditions of success, taking care not to exaggerate them. The ordinary belief does not have to be justified rationally, but it does have to be related to evidence in a way that is commonly accepted as sensible. When co-objectuality is analysed as a philosophical relation, the belief that it holds between two sets of data must be generated by the natural relations included in its analysis, and it must be generated in accordance with principles that we would not reject as foolish or flighty. Of course, the belief may be trimmed a little, in order to get a match with the available evidence, but it must not be shorn of its essentials. Finally, the general principles used in this particular case ought to find their place, and ought to be given some distinguishing characterization, in the general psychological theory.

I do not think that Hume succeeds in this task, and he does not think so either.² His difficulty is clear. When two detached sequences of perceived data are co-objectual, ordinary people postulate unperceived data to fill the gap between them. But he treats the predicate 'unperceived' as a scientific predicate determining a class of data, like visual data or tactual data.³ Now unperceived data are related to perceived data by resemblance when it is a case of constancy, and by causation when it is a case of coherence,⁴ and if the holding of these two relations could be established by experience, they would function as natural relations occurring in the analysis of the

¹ *Ibid.* I. 4. ii.

² *Ibid.*, loc. cit.

³ *Ibid.*, loc. cit.

⁴ *Ibid.*, loc. cit.

philosophical relation, co-objectuality, and explaining the ordinary belief in it. This is how he treats co-objectuality when perception is not interrupted,¹ and if he could treat it in the same way when perception is interrupted, he could explain the ordinary belief in it in much the same way that he explained the ordinary belief in causation. But unfortunately, in this case experience cannot establish that the two natural relations hold,² because some of the related terms belong to a class that is never experienced. This compels him to place the ordinary belief on a weaker basis.

His reconstruction of its foundations is elaborate. What has to be explained is our tendency to extrapolate the constancy or coherence of perceived data beyond the limits of experience. His explanation begins with an analysis of identity. Strictly speaking, he tells us, identity requires both invariance and continuity. But continuity is lacking when all that we have is two detached sequences of perceived data. In spite of this lack, we tend to attribute identity, because our experience of the two detached sequences feels very like our experience of a single continuous sequence. However, we are bothered by the fact that the attribution of identity conflicts with the strict requirement of continuity. So we have developed a second tendency which removes the conflict—the tendency to picture appropriate unperceived data filling the gap between the two perceived sequences. But these pictures, or ideas, need vivacity in order to amount to belief. The crux of the whole account is Hume's explanation of the way in which they get the required vivacity. They get it from the present impressions of sensation belonging to the second sequence, and the present impressions of memory derived from the first sequence. (He often speaks of impressions rather than of ideas of memory.) The relation that transfers the vivacity from these impressions to the ideas of unperceived data is causation. For the impressions cause both the tendency to attribute identity and the tendency to picture unperceived data in order to remove the resulting contradiction.³

If this is how ordinary people come to believe in the existence of unperceived data, their inference is a foolish one. For the causation that functions as a natural relation producing their belief does not relate two sets of observed data as it did in the previous case, the case of the problematical belief in causation itself. What it relates in this case is a set of observed data and

¹ *Treatise* I. 1. vi.

² *Ibid.* I. 4. ii.

³ *Ibid.*, loc. cit.

two mental tendencies. It is obvious that any inference drawn on this principle is altogether flighty.

It is questionable whether Hume's general theory really allows such an inference. As I stated it, the theory requires all operative natural relations to hold between perceived data. This requirement must be met if the idea that some relations have two aspects is going to open up a way out of the dilemma between rationalism and naïve empiricism. That door remains open only so long as the situation can be schematized in the following way: the problematical philosophical relation must be analysed into relations between impressions of sensation, some of which will function as natural relations producing the belief that the philosophical relation really does hold between the impressions of sensation.

This schema amounts to a theory of truth for problematical beliefs. The ideas that make up the belief must form a pattern that matches the pattern of the impressions of sensation. Sometimes, as in the case of causation, it is only mismatch that can be firmly established, and sometimes the pattern of ideas must be trimmed of certain excesses before there can be any hope of a match. But there is one feature of the schema that can never be compromised: any natural relation that serves to produce the association supporting the belief must belong to the reality that is matched, even if it also appears on the signifying side, which bears the onus of matching. The case of causation illustrates this requirement. In that case causation itself, analysed into contiguity, priority, and constant conjunction, relates the original impressions of sensation and that is the essential base of its operations. It then causes on the signifying side an association of ideas, like any other natural relation, and the association itself operates as a causal disposition. But whatever causation does to ideas, it must begin by relating the original impressions of sensation. This condition is violated by Hume's account of the ordinary belief in the continued existence of unperceived objects.

However, when he sets out his general theory, he does not explicitly state this requirement. So it is possible to take the error to lie in the theory itself, rather than in its application to this particular case. Either way, the result is a failure to discriminate between sensible and foolish principles of non-deductive inference.

Hume himself is dissatisfied with his account of the origin of this particular problematical belief, partly because the

mechanism of the inference is so weak. He says, 'I cannot conceive how such trivial qualities of the fancy, conducted by such false suppositions, can ever lead to any solid and rational system.'¹ But he does not locate the central weakness, which is the misplacing of the operative natural relation. He also has another reason for dissatisfaction. He equates the data of ordinary people with impressions of sensation, and he argues that, though it is conceivable that impressions of sensation may exist in isolation, experiment shows that in fact they depend on our organs of sense.² So this particular problematical belief is actually false, and it is to nature's credit that it succeeds in forcing it on us.

I shall not examine Hume's interpretation of the data of ordinary people, or explore other ways out of the impasse into which it leads him. Nor have I time for his brilliant critique of his predecessors' theories of perception. My only point about this section of the *Treatise* is that in it the general theory of the origin of belief is applied in a way that ruins it.

In the Appendix to the *Treatise* Hume does not mention his dissatisfaction with his treatment of the ordinary belief in the continued existence of unperceived objects, perhaps because he assumes that his system could rest on phenomenal foundations. But he does say that his treatment of the third problematical case, personal identity, was mistaken, and that he does not know how to put it right. Since he analyses co-personality in the same way as co-objectuality, it is an interesting question why he is not dissatisfied with the *Treatise* on both scores.

I shall now try to show that he gives the wrong reason for rejecting his earlier account of co-personality. The reason that he gives is that his general theory fails to explain the ordinary belief in co-personality, because it appeals to natural relations that do not support it adequately—more or less the complaint that I have been making against his account of co-objectuality. But the true reason, I shall argue, is something else.

In the text of the *Treatise* co-personality is explicitly modelled on co-objectuality.³ It is represented as a philosophical relation that holds between pairs of mental elements. But it does not hold necessarily between the ideas of any two mental elements, given the contents of those ideas, nor is it based on any special impression. So Hume analyses it, hoping to find that its total legitimate content will include enough natural relations to

¹ *Treatise* I. 4. ii

² *Ibid.*, loc. cit.

³ *Ibid.* I. 4. vi.

explain our belief that it holds. The two natural relations that he turns up are resemblance and causation. Causation produces what he called 'coherence' in his treatment of co-objectuality, but he takes resemblance to produce recurrences in this case, rather than constancies, illustrating it with the example of images of memory. The function of resemblance and causation is to generate associations between the ideas of pairs of mental elements, and so to produce the beliefs that the original elements in the pairs are co-personal.

There is in this case the same difficulty that there was in the case of co-objectuality: he thinks that, strictly speaking, identity is incompatible with change of composition and with interruption, both of which are common in any mental history. Nevertheless, causation and resemblance make the transition from the first member of a related pair to the second member feel like a transition without interruption, and resemblance makes it feel like one without change, and so we are seduced into picturing the mind as a continuing entity. In order to complete his explanation, he has to say what gives such pictures the vivacity required for belief. The main difference between his account of co-personality and his account of co-objectuality is that he is inexplicit on this last point.

This is understandable. For the required vivacity has to be derived from an impression of memory of the earlier element in the pair and from the later element, which, we may assume, is now present to the mind. But the latter may happen to be an idea, in which case, according to Hume's theory, it will not be a source of vivacity. However, it is obvious that in fact a present idea is as good a source of vivacity *for this particular belief* as a present impression. The explanation of this fact is that it is sufficient in this case that there should be consciousness of the idea, as, according to him, there always is. But the theory of vivacity would have to be rewritten to accommodate this point, and that is why he says so little about the source of the vivacity required for the belief in co-personality.

Hume's own reason for his dissatisfaction with his analysis of co-personality and with his account of the origin of the ordinary belief in it is interesting but incomplete. As usual, the account avoids both rationalism and naïve empiricism and relies on a sophisticated criterion of match and on a certain amount of trimming of the ordinary belief in order to achieve it. In the Appendix he is still convinced that this solution is on the right lines, but he says, 'I find myself involved in such a labyrinth

that, I must confess, I neither know how to correct my former opinions, nor how to render them consistent.'¹ His difficulty is that he is unable to 'explain the principles that unite our successive perceptions in our thought or consciousness'.² He has analysed co-personality as a philosophical relation, but he now finds that the two natural relations occurring in its analysis, resemblance and causation, are insufficient to explain the ordinary belief in it. He evidently supposes that there is a shortcoming in the detailed development of his theory, a shortcoming that is shown up when it is applied to the ordinary belief in co-personality, and he is not envisaging the possibility that this belief could not be explained by any variant of his theory.

This is somewhat obscured by his observation that there was some inconsistency in his earlier account of co-personality, especially when he tries to explain what the inconsistency was. 'In short, there are two principles which I cannot render consistent, nor is it in my power to renounce either of them, *viz.* that all our distinct perceptions are distinct existences, and that the mind never perceives any real connection among distinct existences. Did our perceptions either inhere in something simple and individual, or did the mind perceive some real connection among them, there would be no difficulty in the case.'³

It sounds as if he means that the two principles are inconsistent with one another. But he cannot mean anything so absurd. He must mean that the two principles prevent him from adopting a rationalist account of co-personality, and so, since it is equally impossible for him to adopt a naïve empiricist account, he is forced to look for an adequate explanation along the lines of his own sophisticated empiricism. He must find such an explanation, if he is going to compose all the elements—including the essential features of the ordinary belief—into a coherent picture. But, he is saying, he cannot fit all the pieces together, because his philosophy does not allow him to include all the essentials of the ordinary belief. In short, the two principles pose the problem and he is unable to find a solution in which everything falls into place, and the discrepancy between the available evidence and the actual belief is too great to be removed by any trimming of the belief.

However, he does not tell us precisely what is wrong with the explanation offered in the *Treatise*, and he only expresses his conviction that it ought to be possible to find a successful

¹ Appendix to the *Treatise*.

² *Ibid.*, loc. cit.

³ *Ibid.*, loc. cit.

explanation along those lines. I think that the reason why he has no more to say is that he has not succeeded in diagnosing his own earlier error. He believes that what is needed is an improved account of the principles of union that will subsume this difficult case under his general psychological theory. But the truth is that no such account can be given of co-personality.

Since this is a familiar truth today, I shall not dwell on all the details, but I shall end this lecture with a brief demonstration of their impact on Hume's theory.

His diagnosis of his former error is that resemblance and causation do not really explain the union between co-personal mental elements, and that they need to be reinforced by stronger links. But this cannot be right, because such elements do not need to be related in any way through their contents in order to be co-personal. It is true that a sequence that exhibited no such patterns would be a disintegrated mind, but the elements in it would still be co-personal. Of course, the concept of a person would apply to such a case in an attenuated form, because the so-called person would have no sense of his own identity. But such deficiencies do not exclude co-personality.

Hume writes as if he starts with a class of elements directly accessible to himself through present experience or experience-memory, and then has to find a suitably related sub-class, just as he had to find suitably related sub-classes in the other two cases, causation and co-objectuality. But this cannot be right, because in fact nobody has direct memory-like access to others' elements as well as to his own, and so nobody is faced with the question, which directly accessible elements are related to the present moment in his mental history in ways that make them his. Of course, we can imagine the human species endowed with quasi-memory as well as with memory.¹ But Hume is certainly not doing this, but trying to develop a theory on the basis of our present endowments.

When he laments the lack of real connection between co-personal elements, he is overlooking the important fact that one of his could not have been anyone else's. But this fact is not the result of the contents of elements, but, rather, of the sequences in which each happens to begin to exist, whatever its content. To adapt his metaphor, it is not only the actors' parts that begin to exist on stage, but also the actors themselves.

Then ought he to have paid more attention to the relations

¹ See S. Shoemaker, 'Persons and their Pasts', *American Journal of Philosophy*, 1970.

that hold between elements independently of their contents? Certainly, that would be necessary in order to allow for the disorderly contents of a disintegrated mind. But it would not have provided him with another way of showing that his elements are necessarily his. For causal relations between elements that have nothing to do with their contents hold contingently, like any other causal relations, and so cannot be established in the rationalist way, through the contents of the ideas of the related elements. However, if he had investigated those relations that hold between elements contingently and independently of their contents, and if he had set on one side the trivial necessary truth about ownership, he might have called off the search for stronger content-based links. At least, it is clear that its motivation is a mistaken diagnosis of his earlier error.

So the final verdict must be an ironical one. Hume's general psychological theory gives a faulty account of fallible inferences and of the origins of the associations that support them. He himself makes this criticism, but mainly of the application of the theory to co-personality. Yet in that particular case the real trouble is something else.

Throughout this lecture, I have used Hume's terminology, modernizing it as little as possible. In particular, I have expressed his theories in terms of his own distinction between philosophical and natural relations. Because his theory is pluralistic and atomistic, his difficulties are best presented as he presents them, as difficulties about relations. His idea, that some relations have two aspects is the key to the sophisticated theory of match that he develops when he is analysing and explaining our problematical beliefs. Through it we can understand both why the theory fails, and why he thinks that it fails. But it is a noble failure and there is more philosophy in it than in many more successful enterprises.