

Mentalism and the Explanation of Behaviour 4Mentalist Explanations - epistemology and ontology

Before going on to consider Mentalist Explanations of behaviour in relation to the different types of theory and explanation employed by professional psychologists for this purpose, there are two aspects of mentalist explanations themselves which were not discussed in the last lecture. The first concerns what we may call the epistemological aspect of such explanations viz: the nature of the empirical evidence on which the various hypotheses comprising the explanation are based or which can subsequently serve [to] confirm or disconfirm them. The second aspect which we must also consider is the problem of the ontological commitments of such explanations. To what extent are we committed in using such explanations to a belief in the existence or occurrence of mental states and events whose existence or occurrence cannot be determined by objective observation?

The epistemology of mentalist explanation

As in the case of many scientific theories, there are a number of different types of evidence which may be taken, either separately or in combination, as a basis for the postulation and subsequent confirmation or disconfirmation of the hypotheses which comprise a mentalist explanation or prediction of human behaviour. As in scientific theories moreover, the most important type of evidence and the one of most general application, is the history of the successes and failures of the various predictions of outcome which are based upon a given hypothesis. Thus, if most of the predictions as to how an individual will behave which involve the hypotheses that he believes a given proposition *p* are subsequently confirmed, and if in those cases where the prediction is not confirmed there is evidence that the failure of the prediction is due to the falsity of some premises other than the hypothesis in question or to the unpredictable intervention of some chance contingency which prevented the agent from doing what he would otherwise have done, we may be reasonably satisfied that the hypothesis in question is true and one which we can confidently employ in formulating subsequent predictions and explanations of his behaviour.

However if this were the only kind of evidence on which we could base our choice of the very large number of possible alternative hypotheses open to us, an explanatory system with the degree of logical complexity of the mentalist system of behavioural explanation as set out in the diagram distributed in connection with the last lecture, would never get off the ground. For although there is a sense in which Popper (4) is right when he claims that one explicit disconfirmation is sufficient to show that a given explanation is false, what is falsified is the explanation as a whole and not the individual hypotheses which comprise it. Of course, if the explanation as a whole is false, then at least one of the separate hypotheses which comprise it must be false; but it may very well be the case that all the other component hypotheses are true and only this one hypothesis that is false. Furthermore from single disconfirmation such as we are envisaging, we have no means of telling how many of the component hypotheses are false and how many are true or which of them it is that are false and which true. No doubt if we could arrange series of experiments in which we systematically varied the initial cognitive and motivational conditions, we could discover which hypothesis is falsifying the explanation as a whole by a process of elimination. Unfortunately the factors determining what a man believes and wants are so complex that controlled experiments of this type are virtually impossible. Consequently, although the pattern of confirmations and disconfirmations of prediction involving a given hypothesis is an important consideration, it can seldom be decisive unless it is supported by corroborating evidence of other kinds.

Apart from the evidence derived from the confirmation and disconfirmation of predictions involving a given hypothesis, there are five kinds of evidence we can use framing and subsequently confirming or disconfirming our hypotheses as to what a man knows, believes, wants or intends, viz:

- (a) what he says

- (b) what he has been told or has read in some document, book or newspaper
- (c) what he has had an opportunity to observe
- (d) how he has reacted to various contingencies that have arisen from time to time
- (e) by deduction from other independently established hypotheses of a similar kind.

These different kinds of evidence apply in different ways and to different degrees according to the different kinds of hypothesis involved in a mentalist explanation. There are moreover, premisses involved in the prediction of behaviour on a mentalist basis, such as the hypothesis 'conditions are favourable for Φ -ing' or the hypothesis 'A can Φ ', which are based on and confirmed by empirical observation of what other people and the agent have succeeded in doing on similar occasions in the past. But putting those cases on one side, I propose to focus attention on the two main kinds of hypothesis which are peculiar to and distinctive of mentalist explanations, cognitive hypotheses and motivation, in order to see how the different kinds of evidence I have mentioned apply in framing these two types of hypothesis and in testing the various conclusions deduced from them as set out on the diagram distributed in connection with the previous lecture.

Cognitive hypotheses

A cognitive hypothesis may be defined as a component hypothesis in a mentalist explanation of behaviour in which we suppose that the agent either knows or believes a given proposition to be true. As we saw in Lecture 2 and again in Lecture 9-1, to say that a man knows a proposition to be true is to say (a) that he believes it to be true (b) that it is true and (c) that he has good or adequate grounds for this belief. However the question of the truth or falsity of what a man believes only affects our prediction as to how he will behave in so far as he is less likely to encounter disconfirming evidence for his belief, if it is both true and well grounded, than if it is false or held for the wrong reasons. This means that there is much less likelihood of man's changing his beliefs, if they are things he can quite properly be said to know, than if they are merely beliefs which are either possibly or certainly false or, if true, are held for wrong or inadequate reasons. But though it may affect the confidence with which we make the prediction, the question whether a man knows or merely believes a proposition to be true makes no difference to the kind of behaviour that is predicted. Nor is there necessarily any relationship between the truth and falsity of the proposition believed and the confidence with which the agent is prepared to maintain it or the subjective probability he assigns to it. It is for these reasons that I have used the verb 'believe' rather than the verb 'knows' in formulating the sentence frames expressing the cognitive hypotheses on the diagram.

- (a) [Evidence based on what a person says.] We have also seen in Lectures 2 and 9-1 and again in the previous lecture that there is a close conceptual connection between believing a proposition to be true and asserting it or saying that it is true, in that although one may at the cost of lying, assert a proposition without believing it to be true, one cannot assert a proposition without intending that one's audience should believe it; nor can one believe a proposition that one would not be prepared to assert under conditions where one had no occasion or motive to lie or otherwise conceal one's opinion. Indeed a plausible account of what it means to believe a proposition is to say that to believe that p is to be disposed under appropriate conditions (a) to assent to p, if asserted by someone else, (b) to assert p oneself and (c) to act on p i.e.: to act in a manner appropriate to p's being true.

If this account is correct, it follows that if we hear someone asserting or assenting to a given proposition, we have at least prima facie evidence that he believes the proposition in question, evidence moreover, which can only be overturned at the expense of convicting him of mendacity. This then is the principle which we employ in using what a man says as evidence for hypotheses about what he believes.

It should be noted however, that when we fill in our hypotheses about a man's beliefs or about what he knows, we listen to everything he has to say about the question at issue, not merely to those sentences which are expressed in the form 'I know that p' or 'I believe that p'. For these ego-statements are not, as has sometimes been supposed, introspective reports by the speaker on his inner mental state (3), they are simply different ways of asserting the proposition the individual in question claims to know or believe. In many cases the phrases 'I know' and 'I believe' are wholly redundant and can be omitted

without changing the force of what is said; and where they do have a function, they function as what Frege (1) has called ‘a sign of assertion’ indicating that the speaker is genuinely asserting the proposition and not just putting it up for consideration as a sample sentence for logical or grammatical investigation or as a possible thesis about which he has not yet made up his mind. Moreover the distinction between ‘knowing’ and ‘believing’ has a different function when used in the first person to characterise the attitude of the speaker himself than it has when used in the third person to characterise the propositional attitude of someone else. In the third person case ‘He knows that p’ commits the speaker, as well as the person about whom he is speaking, to an assertion of the truth of p, whereas ‘he believes that p’ does not commit the speaker to the truth of the proposition which he attributes to the person he is talking about. However, since both ‘believing’ and ‘knowing’ entail a disposition to assert the truth of the proposition known or believed in the case of the person who is said to know or believe, it follows that a first person statement of the form ‘I believe that p’ commits the speaker to an assertion of the truth of p, just as much as does a first person statement of the form ‘I know that p’. The only difference between the two claims is that the statement ‘I know that p’ involves, in addition to the assertion of p, the claim that the speaker's grounds for such an assertion are such as to leave no doubt about the truth of the proposition he is asserting; whereas ‘I believe that p’ acknowledges that adequate rational grounds for the assertion cannot be supplied and thus leaves open the possibility of his being mistaken. It follows from this that, when used in the first person, the distinction between ‘knowing’ and ‘believing’ serves merely to indicate the speaker's confidence or lack of confidence in the adequacy of his grounds for asserting the proposition he is asserting. To that extent it serves to characterise his mental state in asserting the proposition in question. But this does not alter the fact that the primary function of sentences of the form ‘I know that p’ or ‘I believe that p’ is to assert the proposition p and that other people are perfectly justified in forming conclusions about what a man believes and whether or not he can be said to know it or merely to believe it, regardless of whether or not he prefaces his assertion in this way and regardless of the way in which he himself characterises his attitude to the proposition in question.

- (b) [Evidence based on what a person has been told or has read in some document, book or newspaper.] Listening to what man has to say about the situation in which he finds himself is not the only way of arriving at reliable conclusions about what he knows or believes. We can reach similar conclusions by studying the sensory inputs to which he is or has been exposed. These sensory inputs are of two main kinds. In the one case the prospective agent receives a sensory input which has already been coded in the words of a language which he understands by some other person i.e. he hears or reads a proposition or set of propositions expressed in a sentence or set of sentences which are spoken or have been written, typed or printed on paper by some other person or persons. If, when exposed to a sensory input of this kind, the prospective agent gives no indication that he has either failed to understand or has rejected the propositions in question, and if there are no other reasons for thinking that he would be inclined to doubt the reliability of the informant who has passed on the information to him and hence the truth of the propositions, it is safe to assume that he has come to believe them and will henceforth adopt them as premisses in his practical reasoning.
- (c) [Evidence based on what a person has had an opportunity to observe.] In the second case what we observe is that the prospective agent is being or has been exposed to a set of sensory inputs which have not been pre-coded in linguistic form by someone else, but which we can reliably expect him to code in the form of some sentence expressing a given proposition about the situation to which he has been sensorily exposed and which he will subsequently believe to be true. This is the case where he acquires a belief on the basis of his own observation and sense perception of the situation in question. The psychology of common sense contains no laws or principles, such as the empirical psychology of sense-perception has attempted to construct, which would enable us to predict the conclusion about the situation confronting him which a man will draw and hence what he will subsequently believe from a knowledge of the sensory inputs to which he has been exposed. But since in the vast majority of cases the conclusions which one individual draws on the basis of the sensory input he receives in a given situation will not differ significantly from those which are drawn by other people exposed to similar inputs on the same occasion, we can usually assume that the conclusion which a prospective agent draws

from his observations in a given situation will be the same as we draw ourselves on the basis of our own observations in the same situation.

- (d) [Evidence based on how a person has reacted to various contingencies that have arisen from time to time.] We can also obtain what amounts to confirmatory evidence about a man's beliefs in so far as they relate to future outcomes, i.e. about his expectations as to what is or is not likely to happen under a given set of circumstances, by studying his reactions when he observes occurrences of different kinds. If he displays a marked startle reaction when something happens, this is strong evidence that he did not expect the occurrence in question. If no such startle reaction is evident in his behaviour, this is some evidence that what happened was in accord with what he expected. Unfortunately the evidence for his not expecting a given outcome based on the startle reaction is very much stronger than the evidence that he did expect it provided by its absence, since many people can conceal their surprise, if it suits them to do so. Moreover since the effect of such an experience is to alter the subsequent expectations and beliefs of the individual who is exposed to them, the information obtained in this way has little predictive value for agent's subsequent behaviour.
- (e) [Evidence based on deduction from other independently established hypotheses of a similar kind.] Finally we can draw conclusions about what a man believes by considering the logical consequences of other propositions which we know or suspect him to believe on other grounds. As we have seen, this deduction of the conclusion that a man believes or will come to believe a given proposition by virtue of the fact that it follows logically from other propositions which he is known to believe is one of the most important factors in enabling us to move from hypotheses about what a man says to a prediction as to what he is likely to do. At the same time although we are often able to confirm, either by his subsequent assertion of the conclusion of such an inference, or by his subsequent decision to act in a manner consistent with his drawing such a conclusion that such an inference has been drawn by the agent, we very seldom have direct observational evidence of the occurrence of the thought process whereby the agent deduces the conclusion from these premisses, if indeed, any such explicit thought process actually occurs at all, which in many cases it probably does not. Furthermore we know that there are large individual differences in the case of human beings in the ability to draw the correct inferences from a set of premisses. Many people draw inferences which are not entailed by the premisses of the argument and fail to draw conclusions which are entailed by the premisses.

On the other hand there are some philosophers, Hintikka (2) for example, who are inclined to argue that if someone believes a proposition or a set of propositions, he must also believe any proposition which is entailed by that proposition or set of propositions. For to say that man believes a proposition or set of propositions entails that he is disposed under appropriate circumstances to utter or assent to the proposition or propositions in question. But he cannot utter or assent to a proposition without uttering, hearing or reading a sentence which expresses that proposition; nor can he be said to utter or assent to a proposition unless he understands the indicative sentence in which it is expressed; and he can only be said to understand an indicative sentence in so far as he can draw the correct inferences from the proposition it expresses. From this it would seem to follow that a man cannot be said to believe a proposition if he is not able to draw the correct inferences from it.

However a consideration of relevant examples (e.g. compare the case where a man infers from the premisses 'The train leaves at 12.00 hrs today' and 'it is now 15.00 hrs' the conclusion 'I am in good time to catch the train' with the case where he infers from the premisses 'The time difference between New York and Amsterdam is 6 hrs' and 'The time in Amsterdam is 15.00 hrs' the conclusion 'The time in New York is 21.00 hrs') suggests that it is only in cases, such as the first case where the conclusion contradicts one or both of the premisses from which it is allegedly inferred that we would be forced to conclude that he does not understand the sentences involved, and hence cannot be said to believe all three of them. In the second type of case where he fails to draw the conclusion demanded by the premisses, but where the conclusion he in fact draws is not inconsistent with the premisses, there is no reason either to deny that he believes all three propositions or to assert that he believes the proposition 'the time in New York is either 09.00 hrs or 21.00 hrs' which is the conclusion which in fact follows from the premisses in question.

It follows from this that while we are usually justified in excluding the possibility that a man will infer and thus come to believe a proposition which manifestly contradicts the propositions he believes, assuming that he understands the sentence in which those propositions are expressed, we can never be entirely certain that he will always draw and subsequently believe the conclusions which follow as a matter of logic from those propositions and no other conclusion. Nevertheless the assumption that people usually draw and thus come to believe the more obvious conclusions which follow from the various propositions they believe, provided such conclusions do not contradict other beliefs which they hold, is a working hypothesis whose utility is repeatedly confirmed in practice. However this is another point where empirical studies by empirical psychologists of the relative ease and difficulty of different kinds of inference and the kinds of mistaken inference which are most likely to occur, both in the case of human beings in general and in the case of particular individuals, may ultimately enable us to give greater precision and accuracy to our hypotheses about an individual's beliefs in the context of mentalist explanations and predictions of behaviour.

Motivation Hypotheses

- (a) [Evidence based on what a person says.] When we characterise an agent's behavioural dispositions in terms of the contingencies he wants to promote or prevent, we are not characterising his behavioural propensities in terms of the relationship between what he says or is inclined to say and what he does or is inclined to do in quite as direct a way as we do when we characterise his behavioural dispositions in terms of what he knows or believes. Moreover when a man tells us what he wants or does not want, he is talking about himself in a way in which he is not in any obvious sense talking about himself when he tells us what he knows or believes about the world in which he finds himself. Whereas we learn what a man knows or believes simply by listening to the indicative sentences which he asserts whether or not they are prefaced by phrases like 'I know' or 'I believe', we cannot tell what a man wants or does not want simply by listening to sentences in which he refers to the contingencies which he wants to promote or prevent. We can only fill in our hypotheses about what he wants or does not want by noting those sentences in which he expresses a pro- or con- attitude to the contingency in question by the use of the phrases 'I want' or 'I don't want' or their various synonyms.

Nevertheless it is a mistake to suppose that the primary linguistic function or use of the sentence frames 'I want O' and 'I don't want O' is in making some kind of introspective report on the speaker's current mental state. As Toulmin (7) has pointed out, the primary linguistic function of sentence frames of the form 'I want O' is in choosing or in asking for what one is said to want, while the primary function of sentence frames of the form 'I don't want O' is in refusing or forbidding what one is said not to want. In other words when we want to know what a man wants or does not want we listen to what Skinner (6) calls his 'mands' the commands he gives, the requests he makes and the questions he asks, whereas when we want to know what he believes we listen to what Skinner calls his 'tacts', the statements that he makes.

But whereas as we have seen, we report a man's beliefs by means of the oratio obliqua or reported speech construction, we do not normally use this construction in characterising what he wants or does not want. This grammatical feature moreover, reflect the fact that whereas to believe a proposition is to be disposed to act in a manner consistent with accepting and asserting the statement in the oratio obliqua or its equivalents, to want or not want something is not tied to the making of a certain request or the issuing of a certain command in quite the same way. The point is that to want something is to be disposed to act in such a way as to bring about or at least, not try to prevent the coming about of what one wants, while not to want something is to be disposed to act in such a way as to avoid or prevent what one does not want and not to interfere with anything calculated to avoid or prevent it. Now to ask for something or to order someone to do something is one way of bringing it about, while to refuse or forbid something is one way of trying to prevent or avoid its occurrence. Hence to ask for something or to order someone to do something is to show that one wants whatever it is that one asks for or orders someone to bring about and to refuse or forbid something is to show that one does not want whatever it is that one refuses or forbids. But asking, ordering, refusing and forbidding are not by any means the

only ways of acting so as to bring about or prevent a given contingency and, therefore are not the only ways of showing that one want or does not want something. Any action or any attempt to perform an action which is done with an intention or purpose shows, at the very least that the agent wanted to bring about either the actual or the intended consequences of what he did. Non-verbal actions of this kind moreover, show what an agent wants to bring about or prevent in precisely the same way as do his verbal actions, his askings, orderings, refusings and forbiddings, in marked contrast to the case of beliefs where the assertions that he makes and his acting on the proposition that he asserts provide two quite different and complementary kinds of evidence for the hypothesis that he believes the proposition in question. One of the consequences of this difference between beliefs and wants is that it makes very much better sense to talk of non-language organisms like animals and infants wanting and not wanting things than it does to talk of them having this or that belief. It also helps to explain why we find it much easier to talk of unconscious motives, than we do to talk of unconscious beliefs or unconscious mental processes.

- (b) & (c) [Evidence based on what a person has been told or has read in some document, book or newspaper, or what a person has had an opportunity to observe.] As is implied by the diagram setting out the logical structure of mentalist explanations distributed with last weeks lecture notes, if a man comes to believe that the coming about of a particular state of affairs or the performance of a given action will result in the coming about of a state of affairs which he wants or in the prevention or avoidance of a state of affairs which he does not want, he will necessarily want the state of affairs or to perform the action which he believes will bring this about. Thus by acquiring new beliefs of this kind a man automatically acquires new desires and, pari passu new aversions. Consequently any sensory inputs, whether verbally pre-coded or not, and any inference which he makes which results in his acquiring a new instrumental belief of this kind will result in his acquiring a new desire or aversion, provided that he already wants or does not want the state of affairs which he comes to believe it will bring about. This acquisition of new desires and aversions through the acquisition of new instrumental beliefs on the basis of pre-existing desires for the newly predicted outcome, is the principle method whereby we come to postulate and predict the existence of desires and aversions for which we have no previous behavioural evidence, whether verbal or non-verbal. Where a man wants something for its own sake and not as a means to some other end or where he does not want something, not because he dislikes what he believes to be its probable consequences, but because he finds it intrinsically abhorrent, common sense has no clearly formulated principles for predicting under what circumstances such a desire or aversion will come into existence if it does not already exist.

In recent years however, experimental studies by psychologists of conditioned reinforcement and conditioned emotional responses have thrown some light on the way in which such desires and aversions are acquired and much of the technology of behaviour therapy consists in the attempt to apply these experimentally derived principles to the development of new desires and aversions of this basic kind and the elimination of old ones. It is noteworthy that many of the conditioning situations employed for this purpose closely resemble the sensory input conditions for the acquisition of the instrumental belief that certain desirable or undesirable consequences will follow from a given action or state of affairs. This has no doubt served to reinforce the common sense opinion, based on standard mentalist explanatory procedures, that in order to modify an individual's motivation all that is needed is a modification in his beliefs about the consequences of his actions, ignoring the fact that if the consequences are remote or if the agent's assessment of their probability is low, the strength of the desires or aversions they engender will be relatively weak. Hence the notorious inefficiency of such behaviour modification techniques as preaching and legally ordained punishments in which the aversive consequences are only remotely connected both in time and in terms of observable causal connection with the behaviour which it is intended to modify.

It follows from this that as far as our common sense understanding of matter is concerned there is no effective way of predicting what a man's desires and aversions will be from a knowledge of the sensory inputs he receives, whether verbally coded or not, except indirectly through the effect which such inputs may be expected to have on the agent's instrumental beliefs. But here again we may reasonably expect

an improvement in our ability to formulate hypotheses within a mentalist explanatory framework from empirical psychological research.

- (d) [Evidence based on how a person has reacted to various contingencies that have arisen from time to time.] The study of the sensory inputs which an agent receives can however, throw light on the nature of an individual's desires and aversions in another way which does not depend on the prediction of the acquisition of new instrumental beliefs with a consequent effect on his desires and aversions. We can evaluate existing desires and aversions by studying the way in which the agent reacts to verbally coded and non-verbally coded sensory inputs which are such as to induce beliefs, not about the consequences of certain contingencies and actions, but about what has actually occurred in the agent's environment. In order to understand the basis on which we make such evaluations we need to point out that to say that someone wants something entails, not only that he is disposed to act in such a way as to bring about what he wants, but also that he will be (a) pleased if he comes to believe that what he wants has come about (b) worried or afraid if he comes to believe that it may not come about and (c) either angry or depressed or both if he comes to believe that there is no longer any possibility of its coming about. Similarly, to say that someone does not want something is to say, not only that he is disposed to act in such a way as to prevent the contingency in question from coming about, but also that he will be (a) relieved, if he comes to believe that it is no longer likely to come about (b) worried or afraid if he comes to believe that it is about to come about and (c) angry or depressed if he comes to believe that it has come about. If therefore, we have an opportunity of observing an individual's emotional reactions to either being informed of or himself observing something coming about or failing to come about when expected, we have very strong objective evidence of the nature of what he wants or does not want to come about which does not depend either on observing the requests which he makes or the orders that he gives or the non-verbal actions he performs. What is involved in making such observations, we shall discuss in Section 7 of the course, when we shall come on to consider the problem of emotion.
- (e) [Evidence based on deduction from other independently established hypotheses of a similar kind.] Since inferences only effect what a man wants indirectly by altering his beliefs, nothing needs to be added under this heading to what has already been said either in talking about the way inferences create new beliefs or in talking about the way instrumental beliefs engender new desires and aversions or about the way beliefs about what has happened or is likely to happen elicit emotional reactions which show what a man's desires and aversions are.

The ontological commitments of mentalist explanation

The behaviourist, according to Watson (8) "dropped from his scientific vocabulary all subjective terms such as sensation, perception, image, desire, purpose and even thinking and emotion as they were subjectively defined". I have tried to show in these lectures that at least as far as the key terms 'knowing', 'believing', 'wanting' and 'intending' are concerned the contention that these terms are subjectively defined is without foundation, and that therefore the reason put forward by Watson for dropping such terms from the scientific vocabulary of psychology and with it the whole mentalist system of behavioural explanation which depends upon it is no longer valid. Nevertheless the analysis of this system of explanation does perhaps suggest that it contains ontological implications or commitments which make it unacceptable as a scientific explanation in certain contexts. In particular it may be argued that even if the terms 'behaving', 'wanting' and 'intending' refer, as I have argued following Ryle (5), to dispositions to behave in a variety of publicly observable ways, we are not justified in characterising the behavioural dispositions of an organism, such as an animal or a human infant, which does not speak, in terms of what it is said to believe, since in doing so, we are characterising what the agent does in terms of the relation between what he says and what he does; and such organisms, ex hypothesi do not and cannot say anything.

On the other hand it may be argued that all that we are saying when we say that animal or an infant believes something is that it is disposed to act in the way in which an organism with the power of speech would be disposed to act if it were also disposed to assert the proposition in question. If moreover, there were no alternative way of characterising such an organism's behavioural dispositions there is no reason why we should not characterise an animal's or infant's behavioural dispositions in this way for scientific purposes.

But since one of the most important sources of evidence on which we rely in forming and checking hypotheses about an agent's beliefs, namely what he says is necessarily missing in the case of non-speaking organisms, such hypotheses are necessarily that much more speculative in such cases. Consequently the case for devising an alternative way of characterising the behavioural dispositions of an animal or an infant for scientific purposes is overwhelming, though not for the reasons that Watson gives.

A more serious ontological problem is presented by those parts of a mentalist explanation which involve, not the mental dispositions of believing, wanting and intending which can be plausibly construed as dispositions to behave in publicly observable ways but the mental acts of inferring and deciding and the mental activity of thinking about the problem at issue which necessarily precedes any such act of inferring and deciding. Are we not committed in using these concepts in our explanations of behaviour, to a belief in the occurrence of mental processes which are wholly subjective in the sense that their occurrence can only be observed by the agent himself? The standard behaviourist move at this point, which is made both by behaviourist psychologists like Watson (8) and Skinner (6) and by behaviourist philosophers of mind like Ryle (5) is to point out that although human beings actually conduct their deliberations in such a way that they cannot be detected in so doing by another person, they can equally well and sometimes do, deliberate by talking to themselves out loud in such a way that what they say to themselves can be observed by anyone within earshot. In some cases moreover, the process of reaching a decision or coming to draw an inference may take the form of a verbal interaction between two people in which some of the steps in the argument are supplied not by the agent himself but by another person who is helping him to make up his mind. From this observation it is a short step to the hypothesis that in those cases where a man makes up his mind without any such overt rehearsing of the arguments, either to himself or with another person, there is a similar covert rehearsing of the arguments involving some kind of sub-verbal speech which can be heard or felt only by the agent himself.

We shall have occasion to examine and refine this sub-vocal speech theory of thinking in a later lecture. For the moment suffice it to point out that this theory only allows us to postulate this subvocal rehearsing of arguments in the case of organisms which have the power of speech. This means that we cannot legitimately postulate the occurrence of acts of inferring or deciding in the case of non-speaking organisms like animals and children, nor in the case of those purposive actions of speaking organisms which they are not aware or conscious of doing. However as we saw in Lecture 12, we do not in any case, need to postulate the occurrence of an act of deciding to do something and a consequent preformed intention to do it in accounting for the behaviour of organisms without the power of speech. Moreover, as we saw in the last lecture, we can explain impulsive actions which are not premeditated, as well as actions based on a conscious decision to act, in mentalist terms. Consequently it is only in the case of the generation of new beliefs by inference from pre-existing beliefs that mentalist explanations need to involve ontological commitments which cannot be substantiated in the case of non-speaking organisms or unconscious purposive behaviour. Nevertheless it is not inconceivable even in this case that non-speaking organisms engage in some kind of self stimulating behaviour which is functionally equivalent to the process of drawing an inference in a language-using organism and there may indeed, as many psychologists have maintained be behavioural phenomena in animals and infants which cannot be explained without some such hypothesis as this.

References

1. G. Frege - Philosophical Writings, Trans. P.T.Geach & M.Black, Oxford, Blackwell 1966.
2. J.Hintikka - Knowledge and Belief, Ithaca N.Y., Cornell U.P. 1962
3. U.T.Place - Consciousness and Perception in Psychology, Proceedings of Aristolian Society Supp. Vol.XL, 1966, pp.101-124 esp. pp.120-1
4. K.R. Popper - The Logic of Scientific Discovery London, Hutchinson, 1959
5. G.Ryle - The Concept of Mind, London, Hutchinson 1949
6. B.F.Skinner - Verbal Behaviour, New York, Appleton-Century-Crofts 1957.

7. S.Toulmin - 'Concept Formation in Philosophy and Psychology' in S.Hood ed., Dimensions of Mind New York, Collier 1961, pp.191-203
8. J.B. Watson - Behaviourism, Chicago, Chicago U.P., 1924.