

18. BEHAVIORISM*

I. WHAT IS BEHAVIORISM

I should like to begin by characterizing in a very preliminary way my own conception of behaviorism. Before beginning this characterization, there is something I want to say about the kind of definition I expect to develop. It is philosophically important to be quite clear about the kinds of things or concepts for which it is possible to offer a precise definition and those for which it is not. Consider, for example, the definition of a physical concept like that of mass. It depends on an exact characterization of mechanics or some other branch of physics. On the other hand, the problem is quite different if we are asked to give a definition of physics or of psychology. The source of difficulty in the latter kind of case is that we do not have a well-defined and sufficiently large and flexible formal framework within which we can formulate a definition of physics or of psychology. Of course, it is not simply a problem of the breadth or flexibility of our general framework, but also a problem of the very vagueness and looseness of what we want to mean by physics or psychology. The concept of behaviorism is in many respects much closer to the vague concept of what is physics or what is psychology than to the much more precise concept of mass or of a prime number. For this reason, I shall not attempt in this preliminary discussion of behaviorism to sketch a possible formal definition. I shall, however, have something to say about the characterization of behaviorism as a formal theory.

One initial way to distinguish behaviorism from other approaches to the study of human beings is in terms of the vocabulary used. In behavioristic discussions of human actions or attitudes there continually recur words like 'stimulus', 'response', 'conditioning', 'discrimination', and 'reinforcement'. On the other hand, those who are critical of a behaviorist approach, or those who feel it is not adequate to account for

* This paper has not been previously published. Except for minor revisions it was written during the period 1963-65.

all kinds of human behavior, will emphasize such words as 'intention', 'belief', 'purposive behavior', 'rule-following behavior'. Many of those who like to use these latter terms in a systematic way will favor Brentano's thesis that intentional sentences are required for the description of psychological phenomena, but not for the description of physical phenomena. This formulation of a thesis in terms of sentences is most characteristic of philosophers, but highly unusual for psychologists. An early behavioristic psychologist like John B. Watson, for instance, would scarcely understand the talk about intentional sentences. For him the issue is very clear cut between introspective or subjective psychology of the sort represented by James and Freud, on the one hand, and behavioristic psychology on the other. There has been in the past, though not so much currently, a very considerable literature on behaviorism written by psychologists. Probably none of this literature formulates criteria of behaviorism in terms of the kind of words that occur in sentences, or the kind of sentences that are uttered by scientists. Aversion to talking about sentences is not peculiar to psychologists but is common to scientists of all stripes. For example, physicists who debate the merits of field vs. non-field theories of matter, or who argue about contact-action vs. action-at-a-distance theories of electrodynamics, never formulate the issues in terms of the kind of sentences being uttered. This is true even for such discussions in mathematics, with the exception of those conducted by professional logicians.

Whatever the merits of the linguistic approach to the analysis of behaviorism, it is certainly widespread among philosophers, and there is a very common tendency to convert talk about intentional sentences into talk about intensional sentences. Thus, statements like "John believes that there are lions in Alaska" are not only intentional but also intensional; for the truth conditions of such belief sentences, it is commonly, and I think correctly said, do not satisfy the ordinary extensional truth-functional logic. We mean by this that knowing whether or not there are lions in Alaska does not enable us to determine the truth or falsity of the sentence asserting that John believes that there are lions in Alaska.

The widespread and subtle use of intentional sentences in ordinary talk is not something I see any reason for attempting to exorcise. The task for the behaviorist presumably is to provide an analysis of the truth conditions for such sentences in nonintentional terms. A large and subtle

literature of more than two decades shows clearly enough that this is not a simple or straightforward matter. All the same, I am not at all pessimistic about such an analysis ultimately being given. Later on, I shall attempt to indicate how I think nonintentionally formulated truth conditions for belief sentences can be given. For the present it is only to the point to mention it as a problem.

An apparently very different approach to the issues between the behaviorist and the intentionalist is for the behaviorist to ask the intentionalist at what point in the developing complexity of psychological phenomena, the phenomena become intentional in character. For instance, does the simplest sort of conditioning of a paramecium require intentional concepts for an adequate description? It is interesting to find writers like Chisholm wavering on this point. With the increasing progress of biology and experimental psychology it is surely a difficult thesis to maintain that every kind of conditioning of living organisms requires intentional concepts for their adequate description. On the other hand, if the simplest sorts of conditioning do not require such intentional concepts, it is not easy to say just when intentional concepts do enter. Yet as often occurs with such scientific problems, a case can surely be made by the intentionalist even if he is not able to classify precisely all psychological phenomena into two parts, one of which requires intentional concepts for adequate description and the other of which does not. He can admit that the position of the line which he would draw to make a distinction between the two kinds of phenomena is vague; still he can properly maintain that the concept of intention, and possibly also the concept of the recognition of intention by one organism in relation to another, is central to more complex psychological phenomena and cannot be eliminated or reduced to purely behavioristic terms. It is, of course, part of behaviorism to claim that such a reduction can in fact be made.

The committing of an intentional action and the recognition by another organism that an intentional action has been performed are in many cases closely and intimately related to the use of language. It is another aspect of behaviorism to maintain that linguistic behavior can be analyzed in the same terms as nonlinguistic behavior, without the introduction of any new fundamental or primitive concepts.¹

Although I do not want to attempt to give a formal definition of behaviorism, the discussion of the kind of issues already mentioned can

be pursued much more thoroughly and deeply if a rather definite theoretical commitment about the nature of behaviorism is made. What I propose to do is to formulate a theoretical framework within which the analysis of the issues will be made. I would not claim that this theoretical framework encompasses all that is often meant by behaviorism, or even that it is adequate to the analysis of all problems that constitute central challenges to the behavioristic philosopher or scientist. In the next section I shall give a sketch of the theory, but before turning to that sketch there are certain preliminary distinctions I would like to make.

One distinction treacherously easy to forget is that between making an analysis in terms of a theory and making an analysis in terms of known experimental or empirical facts. Consider, for example, the problem of formulating in extensional behavioristic terms truth conditions for belief sentences. On the one hand, we can attempt this in an informal and intuitive fashion by attempting to describe in a rough way the kind of empirical facts we think can be used to provide the basis for such an analysis. By anecdote, illustration, and reference to the vaguely defined thing called the paradigm case, we can attempt to elucidate how we think such an analysis would go. From a formal standpoint this approach will inevitably be unsatisfactory. An alternative is to attempt to proceed within a well-defined theory. In this case, we would ultimately want an analysis possessing the same kind of formal clarity and rigor that are characteristic of Tarski's definition of truth for mathematical languages. What we gain in precision in this case will perhaps be lost in terms of generality and detailed analysis of particular cases. At this stage a schematic formal characterization is the very most that could be expected, but even in schematic form the formal analysis has the virtue of making clear the weak points as well as the strong points of the proposed behavioristic analysis. In the present case there are, it seems to me, certain difficulties besetting a formal analysis that do not usually arise. If, for example, one presents a formal axiomatization of some branch of physics, the formal properties of the physical concepts occurring in the formal statement of the theory are usually well enough known to permit a reasonable evaluation of whether or not the formalization of the theory is intuitively adequate. It is understood, of course, that the formal statement of theory does not itself make clear how the theory is to be interpreted in relation to experiment. The job of connecting the formal theory to

experiment is itself an elaborate and complicated business requiring a detailed methodological theory in its own right. The peculiar difficulty on this score in the discussion of psychological phenomena, particularly in connection with the use of intentional concepts, is that the intentionalist may claim that the behaviorist's extensionally stated formalism will require intentional concepts in order to provide an adequate interpretation in terms of experiments. I do not think there is any simple way to meet this objection. It can be met adequately only by elaborating a methodological theory of the relation between theory and experiment, and it would take us too far afield to attempt to state in detail such a theory. In the meantime, I am prepared to accept intentional criticisms of my use of theory. (For those interested in how I would begin to formulate the methodological theory to exorcise at this new level the intentionalist ghost, I mention the first two articles reprinted in this volume, 'A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences' and 'Models of Data'.)

My second preliminary point is that the kind of details required of a formal theory of behavior are rather different if we are pursuing particular scientific problems rather than philosophical problems. For the purpose of philosophical analysis many of the particular details of the theory can be omitted; or put another way, detailed formulation of particular axioms about conditioning or responses is not especially relevant to the problems of philosophical analysis. A typical example of a scientific issue, not particularly relevant to the problems of philosophical analysis about behaviorism, is the currently much discussed issue as to whether learning occurs on an all-or-none or incremental basis. From the standpoint of the precise axioms of learning or of conditioning it is a critical matter for many experiments; but it is hard to see that this issue in the psychological literature has much bearing on the philosophical issues posed to behaviorists by the intentionalists. On the other hand, many of the definitions needed for a detailed analysis of intentional acts are not the sort that are of much interest to experimental psychologists. This lack of interest is not because the definitions are too banal or too vague, but because they move in a direction of detail and precision which is either unfamiliar to psychologists or at the present development of scientific work uninteresting to them. The point I am trying to make is that the formulation and use of the theory are rather different when we are dealing with particular

scientific issues than when we are dealing with fundamental philosophical issues. It might be thought by some people that if behaviorism had yet received an adequate theoretical formulation, there would be no problem of the kind I am posing; for there would be one well-defined formulation of the theory, and this theory would be appropriate for all purposes either scientific or philosophical. I wish indeed that this were the case, but as in all areas of currently developing science, the status of theory is not so well defined nor so beautifully etched in detail. It is only for theories long established and now static in character, like classical mechanics or classical electromagnetism, that such an agreed-upon, detailed formulation can be given. Behavioristic psychology is as yet too new a science and at present too rapidly developing to hope to be able to formulate a theory adequate for all purposes. I would like to formulate the theory so that it is general enough to be used with some precision for the analysis of philosophical problems, and yet contains axioms which are in themselves not the sort that will easily be refuted by new experiments performed the day after tomorrow. I know that it is not possible to achieve this objective, but in principle this is what I would hope to do.

II. SKETCH OF FORMAL THEORY OF BEHAVIOR

The key concepts of the theory are those of stimulus, response and reinforcement. The detailed axioms of the theory, which we shall not consider here, are based on the following postulated sketch of the sequence of events on a given learning trial.² First, a certain set of stimuli is presented to the organism. From this set, the organism samples a certain subset. On the basis of the conditioning connections or associations, sampled stimuli and possible responses, a response is made – in the detailed theory it is postulated that the probability of response is simply the proportion of sampled stimuli connected to this response. After a response is made, a reinforcement is given indicating which response was in fact the correct one. On the basis of this reinforcement the sampled stimuli may become reconditioned, that is, a new association between stimuli and responses is set up and the organism enters a new state of conditioning ready for the next trial. Before we go any further, some comment is needed about this talk of trials. Philosophers will be inclined to ask what the discussion of trials has to do with the problems of analysis

confronting us. The answer is that if we attempt a formal characterization of the theory we are naturally led to a characterization closest to that of the majority of experiments performed to support the theory. There is no logical requirement that the concept of conditioning in learning be imbedded in a theory formulated in terms of discrete trials. It is in fact possible to give a continuous-time formulation, but this is a matter of technical rather than conceptual interest. Talk about discrete trials simplifies the job of constructing formal models for the analysis of experiments.

It will be useful to consider an example of the sort of thing the theory ought to cover. We may begin with the familiar Pavlovian conditioning of a hungry dog. This example, as we shall see later, is particularly interesting because Chisholm has claimed it is one of the simplest kinds of psychological phenomena to require intentional concepts for an adequate explanation. In this conditioning experiment the dog comes to salivate when a bell is sounded, and it is common to say that the bell has become the sign of food. One of the things we want to do within our theoretical framework is to offer a formal definition of one stimulus' being a sign of another. To begin with, putting matters in an informal way, I would say that the bell is a sign of food to the dog after a series of learning trials if essentially the following four conditions are satisfied. First, before any learning takes place, the response of salivation is made by the dog in the presence of the food with approximately probability one. Second, before any learning takes place, the probability of the response of salivation taking place upon the presentation of the bell but not the food is approximately zero. Third, there are a series of training trials in which the food and the stimulus bell are presented simultaneously. On these trials the dog responds to the joint stimuli by giving the response of salivation with approximately probability one. Fourth, after the series of training trials, the dog responds by salivating upon presentation of the stimulus bell alone, with approximately probability one.

I shall later want to make a number of comments about this example because we shall use it in the analysis of the necessity of intentional concepts in describing or explaining psychological phenomena. For the present, however, I wish to use it only to suggest a schema for defining the notion of one stimulus' being a sign of another. For the purposes of providing a formal definition I shall use some familiar simple apparatus

from probability theory. We shall suppose that there is an underlying sample space that represents in a formal way the possible outcomes of the experiments we wish to perform. As is customary in probability theory, it is usually not necessary to describe in detail the formal structure of this sample space, but rather to postulate only the probability laws followed by events defined as subsets of the sample space (or, alternatively, random variables defined on the sample space). In the present case the events we shall consider will all have a clear intuitive meaning. I shall use notation that is familiar in the psychological literature. Let R_n be the event of the response of salivation on trial n , let US_n be the event of the appearance of food, the unconditioned stimulus on trial n , let CS_n be the event of the bell's ringing, the conditioned stimulus on trial n ; $\neg US_n$ is the absence of event US on trial n .

With this notation available, we may define CS as a sign of US on trial n . The four conditions in the definiens correspond to the four conditions stated in the particular case of the dog salivating.

- (1) $P(R_1 | US_1) \approx 1,$
- (2) $P(R_1 | CS_1 \& \neg US_1) \approx 0,$
- (3) For $1 < m \leq n, P(CS_m \& US_m) \approx 1 \&$
 $P(R_n | (CS_n \& US_n)) \approx 1,$
- (4) For $n' > n, P(R_{n'} | CS_{n'} \& \neg US_{n'}) \approx 1.$

From the standpoint of philosophical problems confronting behaviorism, there are several general things to be said about this definition and the particular example of canine conditioning, but I shall weave these remarks into the comments on Chisholm's view of intentionalists. However, I do want to note several particular things about this characterization of the salivation experiment. In the first place the definition is so set up that one kind of food familiar to the dog will not be a sign for the other, for according to the second condition, the dog will not salivate in the presence of the sign before the training trials. This would not be true of a different kind of food. Also, conditions (1) and (2) formulate a clear difference between the stimulus and the sign. The fourth condition, on the other hand, excludes most of the stimuli the dog samples in his environment, for he will not salivate upon sampling most of these stimuli.

III. CHISHOLM ON INTENTIONALITY

What I have said thus far about behaviorism will not really satisfy those philosophers who are concerned to maintain the view that much specifically human behavior cannot be adequately expressed or explained in behavioristic terms. The central concept of an intentional or purposeful action is most often cited as an example of a concept that cannot be reduced to the behavioristic notions of stimulus, response, conditioning, and the like. In the remainder of this paper I would like to examine some of the issues surrounding this controversy.

I cannot claim to deal with all aspects of the controversy or even to understand some of them. There are two related but different matters that I feel are particularly relevant. One is the discussion of the impossibility of extensionally or behaviorally defining intentional terms; the other concerns the inappropriateness of causal as opposed to "reason" explanations.

A meticulous and careful defense of the claim that intentional terms used to describe intentional actions cannot be defined extensionally is to be found in the writings of Roderick Chisholm, and it will be sufficient to examine some of his views to express my own attitude to this aspect of the controversy.

Chisholm (1957, pp. 170-171) states three criteria for recognizing intentional sentences. First, a simple declarative sentence is intentional if it uses a substantival expression in such a way that neither the sentence nor its contradictory implies there is anything designated by the substantival expression. For example, "Mr. Bailey hopes to find a three-headed calf to add to his collection of circus animals", "Hilbert wanted to find a decision procedure for the whole of mathematics".

Secondly, a non-compound sentence containing a propositional clause is intentional if neither the sentence nor its contradictory implies that the propositional clause is true or that it is false. For example, "John believes that there are polar bears in Africa", "At one time Hilbert believed that a finitistic consistency proof could be found for the whole of mathematics".

The third criterion is the familiar one that a sentence is intentional when its truth value is disturbed by the substitution of one name or description for another, even though the original phrase and the sub-

stituted phrase designate the same object. Sentences asserting necessity provide familiar examples: "It is necessary that nine is greater than seven. The number of planets is equal to nine. However, it is not necessary that the number of planets is greater than seven".

It is certainly possible to refine the statement of these three criteria, or perhaps to quibble either about their adequacy – whether they cover all cases – or about their classifying as intentional some sentences many people regard as extensional. However, I think they will provide tools of the appropriate degree of precision.

Chisholm then formulates Brentano's thesis that intentional sentences are required for the description of psychological phenomena, but not for the description of physical phenomena. He says the following about the invention of a psychological terminology to describe activities like perceiving in nonintentional sentences.

Instead of saying, for example, that a man *takes* something to be a deer, we could say 'His perceptual environment is deer-inclusive.' But in doing so, we are using technical terms – 'perceptual environment' and 'deer-inclusive' – which, presumably, are not needed for the description of nonpsychological phenomena. And unless we can re-express the deer-sentence once again, this time as a nonintentional sentence containing no such technical terms, what we say about the man and the deer will conform to our present version of Brentano's thesis [p. 173].

Chisholm then goes on to examine three methods of showing that Brentano's thesis is wrong. He first examines the attempt by Ayer and others to describe psychological attitudes in terms of linguistic behavior. Secondly, he examines an approach in terms of the psychological or behavioral concept of "sign behavior". He cites as typical instances of this viewpoint the work of the psychologist Charles Osgood and the philosopher Charles Morris. Thirdly, he examines the attempt to define intentional concepts in terms of the concept of expectation.

There are two sorts of things I want to say about Chisholm's defense of Brentano's thesis. In the first place, I try to show that his criticism of the "theory of sign behavior" is far too simple and crude. But, more importantly, in the second place, I try to argue that Brentano's thesis and similar doctrines about intentionality are essentially irrelevant to the development of behaviorism as a scientific theory.

Concerning nonintentional definitions of 'sign' Chisholm says that such definitions of 'sign' depend upon the substitution of one stimulus for another, or in more standard psychological terminology on the relation

between an unconditioned and a conditioned stimulus. Chisholm focuses his criticisms of this kind of definition of 'sign' on the difficulty of characterizing the respect or degree of similarity between the sign and the stimulus for which it is the substitute. Chisholm summarizes his argument as follows:

Shall we say that V is a sign of R provided that V has *all* the effects which R would have had? If the bell is to have all the effects which the food would have had, then, as Morris notes, the dog must start to eat the bell. Shall we say that V is a sign of R provided that V has the effects which *only* R would have had? If the sign has effects which only the referent can have, then the sign *is* the referent and only food can be a sign of food. The other methods of specifying the degree or respect of similarity required by the substitute-stimulus definition, so far as I can see, have equally unacceptable consequences [pp. 179-180].

Before stating criticisms of Chisholm's analysis, let me say parenthetically that if we were to take his remarks seriously we would be denying the possibility of an objective scientific description of the ubiquitous psychological phenomena of transfer and generalization in learning. The central weakness of his criticism is to present us with essentially only two alternatives: either the sign produces *all* the effects that the original stimulus itself does, or it produces, in some vague fashion, only *some* of the effects. In the first case the sign needs to be the same kind of event as the original stimulus, and in the second case, any two stimuli share some property with the original stimulus. Another primary difficulty of Chisholm's analysis of the Pavlovian example is that he does not take any account of the fact that a sign must be *learned* as a sign of a stimulus. His *analysis* provides no room for changes in the conditioning of the organism.

I have indicated above how an extensional definition of one thing's being a sign for another can be given. The identification of the events R_n , US , CS and $\neg US$ on any trial on which they occur is as extensional as the similar identification of events in physical experiments, and so also is the concept of probability – the same sense of probability can be used and is used for the analysis of psychological and physical experiments.

I realize that my definition is still partially schematic, but I see no essential difficulty in making it as elaborate and detailed as necessary. The conditions for elaboration are not different from those necessary to spell out in satisfactory detail physical or chemical experiments. The point of the constructive definition of being a sign is simply to show that once even a rather meager quantitative apparatus is introduced, partic-

ularly of a probabilistic kind, then Chisholm's criticisms seem to fall very wide of the mark. In connection with the next to last sentence of the quotation from Chisholm, I note that my definition does not require that the signs have effects only the referents can have, and therefore be led into the absurdity that only food can be a sign of food. We turn our attention only to a certain subset of responses of the organism taking place at that time, just as in a physical experiment, for example, one dealing with the behavior of magnets, we examine only some behavior of the magnet and do not concern ourselves with such problems as the shadow cast by the magnet on the laboratory bench, etc. In the salivation experiment, we do not examine the responses that take place in the organism as the food is ingested, nor do we take account of the motions of the mouth in chewing the food, etc. For other purposes it may be desirable to consider such responses, but in every experiment, whether it is an experiment on the behavior of organisms, or the behavior of inanimate things, we are never concerned with all aspects of that behavior, but only with a very selected portion. In this respect, too, it seems that Chisholm's criticisms have gone badly astray.

I turn now to my argument against the relevance, for the development of behaviorism as a scientific theory, of Brentano's thesis or similar arguments in favor of the irreducibility of intentional concepts. I maintain that it is not essential to provide an adequate definition in behavioristic terms of intentional notions in order to develop a quantitative theory of behavior. And this is not because the intentional concepts are wrong or inapplicable to the discussion of behavior, but rather because an introduction of new distinctions and concepts does not require as a necessary prolegomenon the analysis without remainder of the concepts already in the field. The view that such a definitional analysis is needed is based on a kind of completeness claim that cannot be supported in empirical domains. Systematic terms or concepts of any empirical theory are fantastically incomplete or schematic. Whether we look at a theory of learning or a theory of mechanics the point is the same. The concepts of the theory are only loosely connected to any actual experiment. The point I am making is not new; it has been made by many others and I have tried to amplify it elsewhere.³ From the standpoint of sophisticated common sense or the previous theory in the field, the concepts of a new theory may be unbelievably crass and crude in their analysis of the nuances of experience.

This is certainly true of the 17th-century theories of mechanics in relation to their Aristotelian predecessors. It was not taken as a responsibility of these new theories to analyze out in terms of the simple Cartesian and Newtonian concepts the many subtle distinctions introduced by Aristotle and the Scholastics in describing physical phenomenon. Indeed, this would have been a hopeless enterprise. Even so simple a distinction as that of Aristotle's between natural and violent motion would seem to have no clear definition in the Cartesian or Newtonian theories. The Aristotelian doctrines of the potential and the actual and of form and matter had even less chance of being explained.

A small-scale example of such irreducibility is to be found in the development of theories of choice and quantitative theories of belief. One appropriate way of describing modern theories of subjective probability is in terms of the attempt to develop a quantitative theory of belief – partial belief, if you will. It is not feasible to describe here the work that has been done on this subject by Ramsey, de Finetti, Savage, and others. The only point I want to make is that the rather considerable quantitative development of this theory has been made in blithe independence of Brentano's thesis. The problems that still beset the theory do not seem primarily to involve problems of intentionality. Consider, for example, the analysis of the degree of belief in terms of the two-place relation *equal to or less probable than*. If we use the ordinary formalism familiar in probability theory, then this relation is said to hold between two events, although in older writers it would be said to hold between two propositions. I think it is possible to object, on intentional grounds, to the use of the event-language for it may be maintained that the *description* of an event is an important determiner of the degree of belief we assign to an event. My answer to this is twofold. On the one hand, if we are interested in studying beliefs as opposed to the assertion of belief statements, a good case can be made for concentrating mainly on non-verbal behavior. The oft-emphasized rubric is that the true indication of a person's beliefs are the actions he takes and not the statements he makes. Numerous recent writers on subjective probability have emphasized this behavioristic point about measuring degrees of belief. This viewpoint is very well expressed but unfortunately thereafter abandoned in the opening lines of Hare's *The Language of Morals* (1952):

If we were to ask of a person 'What are his moral principles?' the way in which we could

be most sure of a true answer would be by studying what he *did*. He might, it is true, profess in his conversation all sorts of principles, which in his actions he completely disregarded; but it would be when, knowing all the relevant facts of a situation, he was faced with choices or decisions between alternative courses of action, between alternative answers to the question 'What shall I do?', that he would reveal in what principles of conduct he really believed. The reason why actions are in a peculiar way revelatory of moral principles is that the function of moral principles is to guide conduct [p. 1].

In view of these opening lines, it is somewhat surprising to find that almost the entire remainder of Hare's book is devoted to an analysis of the language of morals and not to the development of any theory of moral decisions or actions themselves.

In line with Hare's dictum, the surest way to understand a man's moral principles is to study what he does. Over the past 15 years there have been a large number of experimental studies that are more or less relevant to formal theories of preference and choice (for a review of this experimental literature see Luce and Suppes, 1965). It is characteristic of this literature that it is directly concerned with testing behavioristic theories of choice, and secondly, that problems of intention of the sort raised by Chisholm and other writers do not impinge in any systematic way on the design and execution of the experiments. I do not mean to suggest that this experimental literature makes it a point specifically to deny the correctness or appropriateness of intentional concepts in describing much human behavior. It is rather, as I have already indicated, that these difficult and subtle concepts are bypassed and ignored in the new formal developments of a theory of behavior. As a sample of the kind of empirical findings that are coming out of these studies, let me mention just one that is of some generality. The study of betting behavior both in experimental settings and in real-life settings at the race track show that under a wide variety of circumstances there is a very strong tendency on the part of most people to underestimate high probabilities and to overestimate low probabilities. This kind of finding says something important about the belief structures of average people, and yet obviously does not depend in any way on intentional notions. Perhaps another way of putting a criticism of intentionality theses like those of Chisholm's is that it is not made clear what the relation of the thesis about intentionality is to the scientific study of intentional behavior. The sense in which a scientific theory of human behavior must be intentional for Chisholm is not at all clear.

My second point of emphasis is to stress my conviction that the

peculiar character of intentional contexts to be found in belief statements and in other kinds of modal statements that may be cited in support of Brentano's thesis will disappear once a properly detailed behavioristic analysis of language is given. As a very preliminary indication of how I would conceive doing this, we may consider belief statements about the dog's response to the ringing of the bell. Without quibbling about exact perceptual details, I think we may agree that the event of the bell's ringing in the dog's presence is identical with the event of certain sound waves of a specified range of intensity and frequency reaching the auditory receptor organs of the dog. We do not raise problems about substitutability in asserting, on the one hand, that the dog believes that the ringing of the bell in his presence is a sign that food is to follow, and the parallel assertion that the dog believes that the reception of certain sound waves of a given frequency and intensity range is to be followed by his receiving food. At least, I do not find any difficulties of substitutivity here; and the central reason, I feel sure, is the fact that the dog is not a language user. Yet I would maintain that belief statements about the dog or other non-human mammals are unexceptional. The difficulty with belief statements about human users of complicated language is that a variety of signs are used to encode beliefs, and in a fully detailed analysis it is necessary not only to describe the event about which the belief is held but also the encoding signs. I would defend the thesis that it is impossible to have a belief without such specific encoding. (Whether the organism is explicitly conscious of the encoding is an irrelevant matter.) The essential point is that in terms of the specific encoding one should in principle be able to offer a general definition of truth for belief statements. Thus, if John says, "I believe that there are lions in Alaska" the truth of this statement would be defined in the classical Tarskian manner except that the model is not now a direct model of the real world, but a model of the encoded beliefs of John. It is quite true that the methods for determining whether or not belief statements are true or not will usually be highly indirect, but of course this is true for many statements of other sorts as well. Secondly, the problems of vagueness in belief are not different from the problems of vagueness in ordinary statements. There is no more difficulty in principle in deciding on the truth of a vague statement about belief or a precise statement about vague beliefs than there is in discriminating between the truth of "John ran quickly" and "John ran very

quickly". Perhaps the following analogy will make what I am trying to say about the truth of belief statements clearer. The method of testing the truth of belief statements I am suggesting would be similar to the following test by a computer. The computer has certain information about the world stored in memory. When a sentence is handed to the computer it then applies a Tarskian definition of truth to check the truth of the new statement in terms of the information stored in memory. What is important for the truth of the belief sentence for the computer is the information stored in memory, not correct knowledge about the world. I certainly do not think that what I have said here in this brief way about a nonintentional definition of truth for belief sentences has been sufficiently detailed to solve the many puzzling questions raised about belief statements in the recent literature. I do think it points in the right direction, and in particular, is much closer to the intuitive content of belief sentences than linguistic accounts that involve translating belief sentences in ordinary parlance into statements about belief in certain sentences.

IV. TYPES OF EXPLANATION

I mentioned earlier that an important aspect of controversies surrounding the development of a quantitative theory of human actions concerns the inappropriateness of causal as opposed to "reason" explanations. I had hoped to have time in this paper to devote a fairly detailed effort to refuting the kind of claims that are typified in the arguments given by R. S. Peters in his book on philosophical psychology (1958). Briefly put, Peters' argument is that the rule-following purposive model of human behavior is always required for an adequate explanation of a human action. He admits that causal explanations are relevant and can on occasion state necessary conditions for an action. His argument is that we can never give sufficient conditions in causal terms for a human action, because "we can never specify an action exhaustively in terms of movements of the body or within the body" [p. 12]. As an example of a human action that cannot exhaustively be described in terms of physical movements, Peters mentions the act of signing a contract. He points out the many different ways in which the pen may be held, how the size of the writing or the time taken to finish the signature may vary, etc. His entire position is well summarized in the passage following this example.

So we could never give a sufficient explanation of an action in causal terms because we could never stipulate the movements which would have to count as dependent variables. A precise functional relationship could never be established. Of course, just as we could stipulate a general range of movements necessary to define signing a contract, so also we could lay down certain very general *necessary* conditions. We could, for instance, say that a man could not sign a contract unless he had a brain and nervous system. Such physiological knowledge *might* enable us to predict *bodily movements*. And *if* we had bridging laws to correlate such physiological findings with descriptions of actions we might *indirectly predict* actions. But we would *first* have to grasp concepts connected with action like 'knowing what we are doing' and 'grasp of means to an end'. As such concepts have no application at the level of mere movement, such predictions would not count as sufficient *explanations of actions* [pp. 13-14].

It seems to me that this passage reflects a profound misunderstanding of the nature of scientific method in the physical sciences as well as in the psychological and biological sciences. If we were to take these strictures correctly, no causal explanation in macroscopic physics would be acceptable; in fact, no adequate causal explanation could be given of any physical phenomena at the macroscopic or microscopic level involving motions and interactions of a large number of particles, because we are not now able, and probably never shall be able, to state a precise functional relationship between the motions of the individual particles and the observed macroscopic phenomena. Consider, for example, the thermodynamical and mechanical explanation of the formation of clouds on the windward side of a mountain. This explanation is given in terms of the upward motion of an incredibly large number of air and water vapor molecules. Some general characteristics of this motion can be stated, for example, the mean velocity of a molecule, but it is utterly hopeless to attempt to give any account of "precise functional relationships" between the motions of individual particles and the cloud we can all observe.

Perhaps my deepest objection to what Peters says is that, like Chisholm, he does not seem to recognize the highly schematic character of the causal explanation of any phenomena, animate or inanimate. It is, I would claim, never possible to give a direct characterization of sufficient conditions for the occurrence of a phenomenon. The concept of sufficiency is relative to our description of the phenomenon, and the adequacy of a causal explanation must also be judged relative to that description. There is indeed no such thing as an ultimate causal analysis of any phenomenon. Behaviorism and quantum physics are in the same causal boat afloat on a probabilistic sea.

NOTES

- ¹ For a detailed discussion of this point see the final article in this volume.
- ² See the last article of this volume for a detailed and formal set of axioms.
- ³ Article 2 in this volume.