

A Study of Radical Behaviorism

THOSE who cry "No politics" often thereby support bad politics, and those in whose prefaces philosophy is abjured often proceed to expound bad philosophy. In this paper, I want to examine the views of a man who has recommended the abolition of theories: Professor B. F. Skinner.

But I shall not be trying to show that Skinner's theories are bad; I wish to show only that he does employ them, and that his general arguments against the adoption of theories (or at least certain kinds of theory) are not altogether satisfactory. It is, indeed, only because I regard his work so highly and his arguments as so persuasive that I hope to compensate a little for his influence in the methodological sphere. I shall be especially concerned with his 1954 article "Critique of Psychoanalytic Concepts and Theories" (CPCT) in the *Scientific Monthly* (10; reprinted in this volume), but also with his amplification of certain of these points in his 1953 book *Science and Human Behavior* (SHB; 9) and in the 1950 article "Are Theories of Learning Necessary?" (ATLN; 7).*

The general point I hope to make is this: Skinner's position on almost every issue admits of two interpretations—one of them exciting, controversial, and practically indefensible; the other moderately interesting, rather widely accepted, and very plausible—and Skinner's views quite often appear to be stated in the first form but defended in the second.

Specifically, I wish to suggest that

1. Skinner's idea of the relationship between the "pure" molar behavior approach and neurological, mentalistic, or conceptual theories

* The present paper developed out of an attempt to improve the comments which I made as a member of the symposium at which CPCT was first read—in Boston, December 1953—and which were published in the same issue as that paper. Three very good discussions of Skinner's earlier work should be noted: at a general level, by Dr. Feigl in his paper "Principles and Problems of Theory Construction in Psychology" (PPTCP; 1); with reference to ATLN, by Arthur Ginsberg in "Operational

is seriously oversimplified in a way that, if corrected, renders untenable most of his objections to the latter.

2. Skinner's analysis of causation, explanation, and classification is subject to correction in the same way.

Imagine that, as a trained psychologist, you are a member of the first survey team sent out to the Alpha Centauri system, and that an inhabited planet is discovered there. On landing, you find that the atmosphere is unbreathable and that you have to remain in a space suit. The first sight you get of the inhabitants raises the question whether they are actually indigenous, since they also appear to be wearing space suits. You might think of the alternative possibilities that severe changes in the atmosphere have required them to devise these garments in order to survive, or that they come from a more hospitable part of the planet, perhaps underground. These hypotheses are not theories or parts of theories in Skinner's sense. Suppose that, after a long period of study, you discover a large number of laws concerning their habits, needs, and social structuring, and you organize these into various categories so that you can predict and control and understand the behavior of the Centaurians, (who exhibit, in particular, a remarkable power of memory and a great fondness for aniseed); moreover, you can converse with them and exchange technological data with them. Thus fortified, you consider publishing some of this material in a paper called "A Theory of Social Behavior Among the Centaurians." Skinner would prefer that you refrain from using that title, because he would not wish you to imagine that he was opposing that sort of an activity. But suppose that one day it occurred to you that the extraordinary success exhibited by the Centaurians in recognizing people apparently on the basis of descriptions alone might be due to small television signal receivers in their suits feeding direct into their optic nerves: then you have a theory ("any explanation of an observed fact which appeals to events . . . at some other level of observation" [ATLN, p. 193]). As a matter of fact, even a complex mathematical formalization of your data would count as a theory, at least in a sense (ATLN, p. 216); and certainly, if you introduced such concepts as "excitatory potential of a habit" in organizing your thought about Centaurian dispositions, let alone "the significance of the Terran

Definitions and Theories" (ODT; 3); and, with special reference to the 1938 book *The Behavior of Organisms* (8), by W. S. Verplanck in his 1954 paper in *Modern Learning Theory* (11).

landing in the eyes of the Centaurian Council," you would indeed be guilty of theorizing (ATLN, p. 195; SHB, p. 89).

Now Skinner has two approaches to theories in his sense of "theories." At the very least, he wants to argue that bothering with them is pragmatically unjustifiable (ATLN, p. 194). This is, of course, not the same as saying that such theories do not produce practically useful results: it is to say that there is a *better* way to obtain these results. *In view of this better way*, theories are, pragmatically speaking, a luxury. Professor Ginsberg cannot, therefore, get very far by calling on Skinner "to account for [the] fabulous success [of micro-theories] in physics" (ODT, p. 241) since Skinner would not have to deny this success, although he would, at the very least, suggest that other methods might have been preferable. ("It would be foolhardy to deny the achievements of theories of this sort in the history of science. The question of whether they are necessary, however, has other implications . . ." [ATLN, p. 194].)

But Skinner will, on occasion, go considerably further than this. He regards it as distinctly possible that the attempt to theorize has resulted in a net absolute loss in the history of science—not a relative loss compared to the alternative he prefers, but an over-all loss. It would appear that this position could be seriously supported only if we can establish that there are more unsolved problems or pseudo solutions in science today than there have been successes. Now it is clear that in terms of Skinner's favorite criterion, the possibility of control, we have advanced enormously since, say, 1300. How can we evaluate the losses? Certainly, more scientists are working on more problems than ever before, and in that very positivistic sense the area of darkness is larger; but even under Skinner's direction the same situation would of course arise, so this will not support the relative point he makes, let alone the absolute one. If we are not going to appeal to the number of problems known to be unsolved for evidence of retrogression, then we must insist that many of the solutions in which we commonly believe are, in fact, unsound. Clearly, this alternative is more consistent with Skinner's general position. Throughout his work runs the trail of the debunker, the cry of a man who is by no means afraid to say that what we have always known to be true is, in fact, false—for example, our belief in what he calls "the traditional fiction of the mental life" (CPCIT, p. 302). Could he really point out enough pseudo explanations to show that the progress of modern science has been apparent, not real? To do so, he would

certainly have to pass very adverse judgment on many fields other than his own. I think we should conclude that, on the whole, when adopting this extreme position, he is trying to bring us up short, make us worry whether the approach we have taken for granted can really be justified rather than trying to convince us of a literal truth. One of the reasons why it would be so difficult to show that there has been a net absolute loss from the use of theories in the history of science is the curious disanalogy between net progress in science and net profit on a balance sheet. If at time t_2 we definitely have certain laws which predict with some reliability and which we did not have at t_1 ($t_1 < t_2$), and we have not lost any laws which we did have at t_1 , then there is a straightforward sense in which it can be definitely said that our science has progressed—no matter whether we have also adopted large numbers of useless theories or misleading explanations. We cannot regard an inaccurate belief as cancelling out an accurate one in the way that a loss cancels out a profit. Of course, it is certainly possible that progress has been slowed up by such paraphernalia, but this is a comparative point, a more moderate point. *Relative to other methods*, we may have done poorly. We shall certainly consider this interpretation of Skinner's position.

But there is a more subtle point involved, one which does not fall naturally into the two categories we have established—reasons supporting the relative loss position and those supporting the absolute loss position. While Skinner would not be able to support a claim of actual error in other sciences, he would—as I understand him—be more inclined to argue that the theoretical approach has stultified or prevented research into certain important and actually open questions. These questions concern the correlation between the supposed theoretical event or state that is said by the theorist to explain the observed phenomena (in our case, the minute television receiver is said to explain the hyper-recognition phenomenon) and its own antecedent causes (in our case, the action of the Centaurs when they inserted the apparatus). According to Skinner, the basic fallacy of the theoretical approach is the idea that a proper explanation of the observed phenomena can be given in terms of the theoretical event alone. For Skinner, the essential question would still remain, What are the causes of the theoretical event? Referring to learning theory, he says, "When we assert that an animal acts in a given way because it expects to receive food, then what

began as the task of accounting for learned behavior becomes the task of accounting for expectancy" (ATLN, p. 194). My justification for believing that Skinner at least suspects that this criticism applies to science in general stems not only from hearing him say it but also from the fact that he always avoids denying it in print, even when it would greatly simplify his argument, and that he regards it as closely analogous to metaphorical thinking, of which he says ". . . if modern science is still occasionally metaphorical, we must remember that theory-wise it is also still in trouble" (CPCT, p. 101).

Yet, would he really deny that the discovery of a television receiver in the Centaurian cranial enclosure constitutes an adequate explanation of the observed phenomenon? In an attempt to answer this, we may ask what he says about the corresponding explanations in the field of human behavior. He is willing to concede (in CPCT) that a neurophysiological account of neurosis may one day provide a useful accessory to, or translation of, a science of behavior (p. 302). But the tone of ATLN is less compromising; although he does not wish to deny that neural events may "actually occur or be studied by appropriate sciences" (p. 194), he does not concede that they will ever be in the least useful for a science of behavior—and the same goes for conceptual and mentalistic theories. In SHB, he again makes clear that he thinks neurology, at best, a different subject (p. 28) and, at worst, a misleading fantasy when he suggests that shock treatment or medical treatment of psychiatric disorders is based on a conception "not far removed from the view—which large numbers of people still hold—that neurotic behavior arises because the Devil or some other intruding personality is in temporary 'possession' of the body" (p. 374). Here, again, we have an example of one of Skinner's extreme positions emerging for a moment—a position which, in the three years since SHB was written, has become much less tenable. As he himself has said, "Advance estimates of the limits of science have generally proved inaccurate" (SHB, p. 20). But Skinner's extreme position in this issue is not merely a matter of pessimism about certain lines of research. It springs from a methodological point. He is not merely arguing that reference to theoretical, e.g. neural, events will stultify our thinking, "create a false sense of security, an unwarranted satisfaction with the status quo" (ATLN, p. 194); he wants to suggest that it is, in some sense, a logical error to suppose that such an analysis can solve our problems in a science of behavior.

Eventually a science of the nervous system based upon direct observation rather than inference will describe the neural states and events which immediately precede, say, the response, "No, thank you." These events in turn will be found to be preceded by other neurological events, and these in turn by others. This series will lead us back to events outside the nervous system and, eventually, outside the organism. (SHB, p. 28)

Thus the basic question is only postponed: What are the relations between these original independent environmental variables and the eventual behavior? Neurology has not really advanced the inquiry.

It is worth pausing for a moment in our survey to note that we have already discovered an extreme and a moderate interpretation of Skinner's position on (1) the role of neural theories in a science of behavior (either methodological mistakes, or merely pragmatically worthless); (2) the extent of the role of theories in various fields (unnecessary in any science; unnecessary in a science of behavior); and (3) the success of theories (more harm than good; more harm than the functional approach). The combinations of these points alone provide him with eight alternatives, and I make no excuse for trying to simplify the issue by making some judgments as to the foundation-stones of his methodological analysis and concentrating on these points.

Let us begin by distilling out of ATLN Skinner's actual answer to the rhetorical question comprising the title, but using the term "theory" in its ordinary sense. I take Skinner to be saying, "Theories that are very abstract are premature; theories involving neural, mentalistic, or postulationally defined concepts are (a) misleading, (b) possibly based on a logical error, and (c) certainly unnecessary, since theories involving terms defined operationally by reference to observable behavior are perfectly adequate."

I am not interested in the merely verbal question whether we attach the label of "theory" to Skinner's position, but I am interested in the underlying analyses which have led him to reject it, and me to favor it. There must be, if there is any logic to his position, certain reasons why Skinner thinks that the gap between his methods of prediction and control and those of Hull is more significant than the gap between Hull and, say, Allport. I wish to examine those reasons a little further and see whether we cannot come to some conclusion about their validity and the validity of their extrapolations as criticisms of psychoanalytic theory in particular.

Certainly, we can agree that Skinner's position is less theoretical than even Tolman's; the question is whether it is non-theoretical. I shall argue that it is not, and that Skinner has elevated the relatively untheoretical nature of his approach into a sterile purity that his approach fortunately lacks, thereby illicitly obtaining the semantically somewhat shocking slogan, "The science of human behavior requires no theories."

"Theory" is ordinarily taken in a wider sense than Skinner's. Webster gives as the two most relevant senses, "A more or less plausible or scientifically acceptable general principle offered to explain phenomena" and "The general or abstract principles of any body of facts"; and defines "abstract" as "considered apart from any relation to a particular object." An abstract principle is, hence, one which involves concepts that refer to a class rather than to an individual. Ginsberg starts with a very similar definition (ODT, p. 233), but later goes on to tighten the notion up in a way that is clearly related to a particular philosophical analysis (originally that of N. Campbell), although it, too, would cover Skinner's account of behavior. Feigl has argued (especially in his article "Some Remarks on the Meaning of Scientific Explanation") that a distinguishing characteristic of the set of propositions comprising a theory *T* of domain *D* is that we can deduce ("explain") all the empirical laws *L* of *D* from any proper *T* of *D*; whereas from any *L*, he says, we can only deduce particular facts *F* of *D* or other *L*'s of *D*. Thus $(T \rightarrow L \rightarrow F)D$, where " $\rightarrow$ " represents entailment. Skinner's position would then be expressed by saying that there is no need to go beyond the *L* level, except for purposes of guessing at further possible *L*'s—and such guessing has no place in a scientific report, which should contain only the results of the tests performed on the putative *L*'s. The appropriate reply is that this procedure does not answer the perfectly proper question, "Why do these particular *L*'s hold in this particular *D*?" Skinner's response could take several forms:

a. The question is not proper because it is not within the province of science to ask why the world is the way it is, only to describe how it is. But Chapter 3 of SHB is entitled "Why Organisms Behave," so Skinner does not reject such questions; however, the answers he offers do not apply to the basic laws—e.g., the Law of Effect in his system. Clearly, it would be unsatisfactory to reject such a question in connection with any law (unless an illegitimate type of teleological answer is expected); we can give a perfectly respectable scientific answer to the

question, Why do rarefied gases at normal temperatures obey Boyle's Law? in terms of the kinetic theory and gas molecules.

b. The question is not proper *within a science of behavior*. If a science of behavior is concerned to "clarify these uniformities [the ones noticed by everyone in the course of their ordinary social contacts] and make them explicit" (SHB, p. 16), then its task does *not* extend to explaining the basic laws themselves. Yet it is quite obviously appropriate for some part of the scientific enterprise to be devoted to such questions, and they do concern behavior. This answer is, therefore, in part, a persuasive definition of the term "science of behavior"; it is as though I were to say, "Chemistry is concerned to describe the composition and the interaction properties of all substances." Then a man who tried to find out why an acid combines with a base to produce a salt would not be doing chemistry. Well, we could call him a physicist or a physical chemist instead, but there's nothing illegitimate about his activity; and we may reasonably expect his discoveries to be very useful to chemists, e.g., in predicting ways in which substances with only *some* of the properties of an acid may be expected to react with bases. And here too, of course, it *might* be possible to answer this question without recourse to the micro-level.

c. Skinner could simply reply that *nobody* knows why these particular *L*'s hold yet, so the only practical way to do a science of behavior is to find out more *L*'s. For example, he says, "At the moment we have no way of directly altering neural processes . . ."; hence, reference to neurophysiology "is useless in the control of behavior" (SHB, p. 34). (Although this section begins by dealing with *explanation* rather than *control*, he does not make any comments about the possibility that neural processes might have explanatory power even if we could not manipulate them; compare the possibility of explaining death as due to an inoperable cancer.) Again, he says, "The objection to inner states is not that they do not exist, but that they are not relevant in a functional analysis. We cannot account for the behavior of a system while staying wholly inside it . . ." (SHB, p. 35). This is a somewhat stronger position than saying we can't do it yet, though not as strong as saying that inner states are completely irrelevant (position b.), since, at the end of the same paragraph he says, "Valid information about [inner states] . . . may throw light upon this relationship [between antecedent environmental conditions and eventual behavior] but can in no way alter

it." Superfluous certainly—"we may avoid many tiresome and exhausting digressions" by examining the relationship directly, without reference to inner states—but not quite irrelevant.

The only difficulty inherent in Skinner's position with respect to this point, then, is that he cannot offer an explanation of his basic laws within what he regards as the limits of a science of behavior, and he can offer only one rather restricted type of explanation of the other laws, although he suggests that the functional approach can provide a complete account of behavior. It is most important to realize that he does provide an explanation of many observed laws of behavior—by deducing them from more basic laws together with some particular information—and it is partly for this reason that I think it appropriate to say that Skinner has a theory of behavior. For, although theories of domain D often do refer to events that are micro-with-respect-to D, and I would go so far as to say that it is almost always the case that there is, or should be, one or more such T for any D, there are some T, in the ordinary sense, which are exceptions. The theory of evolution, for example, seeks to explain the origin of species, amongst other phenomena, in terms of observable small differences between individuals of the same species and the differential survival value of these variations in a complex environment. The theory of gravitation, with the basic law formulated in terms of acceleration rather than force, is another possible candidate.

Looking at the Webster definition again, we can see that Feigl might be inclined to feel it included too much, even mere laws, since there is a sense in which they could fall under the heading of "general or abstract principles of any body of facts" and could each be regarded individually as a "general principle offered to explain phenomena." But the law by itself would not, I think, be regarded as providing the whole set of "abstract principles" of D nor, usually, as providing a complete "explanation" of D. A matter of emphasis, perhaps, and the theory of gravitation is little more than a law. It is chiefly the existence of a selected ordered set of basic laws, carefully chosen so that the entire range of empirically observed laws can be derived, that turns Skinner's system into a theory; though, as I shall argue below, other important aspects of it would suffice, especially the terminology. Yet I have used the qualification "usually" above, because there are certain borderline cases in which we talk of a theory without undertaking to go further

than (though we do more than) stating the law, e.g., the "theory of gravitation." I think there are certain special reasons why we do not regard it as necessary in these cases to go beyond the law (cf. Skinner's inability to derive the Law of Effect itself), and I would regard them as showing that Feigl's criterion is too narrowly drawn. For reasons that will be expounded in the next volume of this series, I think the basis for Feigl's criterion can be shown to be a comparatively unimportant formal truth; and its general, but not inevitable, applicability is due to an empirical fact, namely, that we very rarely get to the point of isolating a particular D from other D's without having discovered some L's of D. That we sometimes do is illustrated by the cases above.

But the fact remains, and we shall need to remember it below, that Skinner's system offers no explanation of its basic laws. Describing a demonstration of the Law of Effect, Skinner says, "Anyone who has seen such a demonstration knows that the Law of Effect is no theory" (ATLN, p. 200). Anyone who has seen an Eötvös balance give a direct reading of the gravitational pull of a block of lead might be tempted to say something similar. But, just as the law of gravitation forms the foundation of the theory of gravitation, the Law of Effect forms the foundation of the theory of behavior. It is even more obvious that a system based explicitly on probability of response ("The business of a science of behavior is to evaluate this probability and explore the conditions that determine it" [ATLN, p. 198]) cannot explain every response individually with more than a probability determinable only by prior manifest behavior. There is a limit to the resolving power of the molar approach, and Skinner takes it to be a certain level of probability in explaining the individual response. It is a perfectly respectable scientific enterprise to examine the proximate (neural) causes of a response in order to increase this probability, whether or not we can do it now or prefer never to call it part of a science of behavior.

In understanding the role of theory in psychology, the most important single point is that in this subject, unlike every other scientific subject as it now stands, a considerable proportion of the highly inferential knowledge (or supposed knowledge) is embodied in our everyday talk; and there is a great deal there. It took many thousands of years after man devised a language for him to produce a satisfactory science of dynamics; but common sense had long since incorporated a "theory" of motion according to which certain norms were defined

such that only deviations from these norms required explanation, and a very comprehensive pattern of explanations was available. Perhaps we could call it an early application of "pre-Aristotelian dynamics" when Egyptian bird-hunters blamed the wind for their inability to cast a throwing stick the usual distance, but we might as well call it common sense—a common-sense theory. Today, the theory of dynamics is very far removed from the everyday language. In the same sense, the language of "possession" in the history of psychopathology and that of "humours" in psychology were as much part of the theoretical activity as appeal to "reinforcement" or "reserve strength" have been recently. The difference is that both are far less abstract than terms from current dynamics. The mistake is to imagine that they are not past the logical threshold for the application of the term "theory."

To much of this, Skinner would agree. ("We all know thousands of facts about behavior. . . . We may make plausible generalizations about the conduct of people in general. But very few of these will survive careful analysis. . . . The next step is the discovery of some sort of *uniformity*. . . . Any plausible guess about what a friend will do or say in a given circumstance is a prediction based upon some such uniformity" [SHB, pp. 14–16].) I am more struck than Skinner is, apparently, by the extent to which we successfully predict the behavior of our friends. Now, the language which reflects these uniformities is the sophisticated language of character traits, of attitudes, needs, beliefs, of degrees of intelligence and carelessness. It is not the naive observation language of acts, colors, sounds, shapes, and sensations. Whether one calls this vast conceptual apparatus a theory or merely the common-sense language of behavior analysis is unimportant. It is important that this language is an enormously long way from the naive observation language: it tells us a great deal about the probabilities of certain responses, though not with the accuracy we would like; if it did not, indeed "we could scarcely be effective in dealing with human affairs" (SHB, p. 16). What militates against calling it a theory in a scientific sense is its lack of organization and of explicit formulation, the absence of a sufficient number of basic experiments whose results can be predicted with extreme accuracy and of knowledge about the results of cross-validation checks, inter-judge reliability, and so on. It is precisely these objections that Hull and Allport were seeking to overcome and which make us all, including Skinner, willing to call their systems

"theories." But it is also these remaining (scientific) imperfections in the common-sense language about behavior which Skinner himself has sought to overcome by the careful choice of carefully defined categories, by the careful organization of the laws which he has discovered, and by the standards of explanation that he attempts to meet. Let us now examine the nature of the terms, as opposed to the laws, of theories.

There are two dimensions to the dichotomy "theoretical term-observational term." The one Skinner is conscious of is that along which the appropriate dimensions of measurement change, means of or possibility of observation alters, and technique of definition varies (from postulational "implicit definition" to "ostensive definition," presumably via operational definition as he sees it). But the other dimension is just as important. This we might call the teleological as opposed to the ontological dimension, although, of course, both are methodological in the general sense. Whereas one can invent or select one's terms (i.e., entities) according to either of the two positions Skinner would distinguish as the theoretical and the functional points of view—i.e., according to one's ontological assumptions, *presupposing the purpose of explanation-prediction*—so along the other axis one introduces or adopts one's terms according to different purposes, ranging from the sense-datum or visual-field neutral-description vocabulary to the explanation-prediction, i.e., the theoretical language, presupposing some ontological position. And, on this axis, Skinner is fairly and squarely a theorist. He selects the segments of behavior to label as a result of theoretical considerations, i.e., considerations of predictive power, experimental verifiability, etc. Naturally, the psychology of learning being in the state of development described above, where even the most highly theoretical terms of any "plausible or scientifically acceptable" account are comparatively close to common-sense language, we can often just as well talk of these considerations as being those of common sense rather than of theory. But common sense is what makes us obey speed limits, at least when a patrol car is behind us, and what leads us to fill our gas tanks at the second or third "Last Stop" before crossing the Nevada desert. These are not very theoretical activities—i.e., they involve very rudimentary induction and merely hedonistic evaluation. The particular type of common sense we need in picking out the categories for a science of behavior incorporates, according to Skinner, the belief that we should so choose them as to obtain data "showing orderly changes characteristic

of the learning process" (ATLN, p. 215). But, *ex hypothesi*, we have no reliable account of the learning process at this point. What makes Skinner think that the changes "characteristic of the learning process" are really "orderly"? This is an excellent example of a theoretical conclusion already embedded in the common-sense language. We can call it a theorem of the background theory, to use Professor Wilfrid Sellars's term, which guides our continuing attempts to provide a more powerful and further-reaching theory. Skinner fully recognizes the existence of this background theory—except that he does not apply the wicked term itself—both in the above quotation and later, when he instructs the psychologist to discover his independent variables "through a common sense exploration of the field" (ATLN, p. 215; *italics mine*). For the reasons given, it seems sensible to recognize that this is a very special part of common sense—that part concerned with the explanation-prediction-control approach to the field of learning—and we can better call it a background-theoretical "exploration of the field." After all, Skinner would hardly wish to insist that common sense could not be radically mistaken.

Let us look more carefully at Skinner's practice and see if we can detect other cases where the background theory is affecting his choice of, or definition of, terms, remembering that "Beyond the collection of uniform relationships lies the need for a formal representation of the data reduced to a minimum number of terms"—as long as it does not "refer to another dimensional system" (ATLN, pp. 215–16).

Skinner has always had one rather imprecise notion in his vocabulary—one which occupies a key position: the idea of *topographically differentiable* responses. The way in which this term is treated is characteristic of the interplay between background theory and the current working vocabulary of the developing theory. We begin with the observation that certain instrumental or motor activities are associated with certain objects in the world: we pick up and chew bananas; we throw stones; we embrace our parents; etc. We wish to produce a general account of the development of all these associations, since we have observed certain similarities in the developmental process. We want our generalization to extend, not only over two successive movements of the jaw (that would be the subject for another study, e.g., on extinction), but also over all these activities. Now the most promising looking lever in our background theory derives from our longitudinal S-R studies in

everyday life, and we wish to retain a link with them; so the word "response" is appropriate. The sense in which the various responses mentioned above are different from each other appears pretty straightforward, so we talk of "topographically different responses." Then the troubles begin: Is the response of raising the hand "topographically different" from that of raising the arm? Background knowledge makes it obvious that these two could not be conditioned wholly independently; indeed, it leads us to believe that there will be certain types of relationship between the times required to successfully condition first the hand and then the arm. How are we to cut the pie of past and future experiments? Vast reservoirs of evidence to aid our choice are already in the history of scientific theory-building. We select as the basic term "topographically wholly independent responses" and include "induction," "reinforcement," and "latency" as three other useful-looking bricks with which to construct our account. Notice the powerful background-theoretical content that has been built into each of those extra bricks by their previous existences in other theory-structures, including the background theory, i.e., the common-sense explanation systems. This makes them far more useful, though, of course, when we have begun to see exactly how this building will have to be shaped, they may have their dangers for the new bricklayer who has not yet caught up on their latest connotations; he may tend to treat them in a previously but no longer appropriate way. Thus the word "response," which began in Pavlov's day as a fairly simple-to-use brick, in Skinner's system has to cover not just "striking-a-key" but "changing-to-the-other-key-and-striking" as opposed to "striking-the-same-key-a-second-time" (ATLN, p. 211); and it begins to become less readily distinguishable from what previously would have been identified as a series of responses. But if we regarded it in that way, we would find it quite impossible to uncover those very desirable "orderly changes characteristic of the learning process" which the background theory very much prefers to have at the expense of a little redefinition of the term "response." The operant versus respondent distinction can be woven into the system here to tighten up the incipient fraying.

Meehl and MacCorquodale provide (5, pp. 218–31) an extremely thorough analysis of the maze through which any psychological theorist must guide his definition of the term "response." It is clear, as they point out, that defining the response in terms of achievement is a very

easy way to build one's theory into one's definitions with deplorable consequences for the interpretation of the experiments said to be confirmatory of the theory. They recommend the elimination, in the definition of "response," of "language which refers even implicitly to the properties of other intervals or to stimulation not present to the organism at the time the response is being admitted" (p. 231). This is an excellent prophylactic with respect to the disease of defining responses in terms of goals. But I am not sure that, in the process, it does not make life awkward for Skinner and others who refer to the past in defining some of their responses, e.g., changing keys. In such a simple case, perhaps we can suppose stimulation to be still present from the past behavior without trouble. But more complex cases referring to organisms responding and reinforced at a very low rate—e.g., to visual stimuli by differential blinking—would render progressively more theoretical the assertion that the response of "changing response" refers only to stimulation present at the time the response is completed * ("emitted" will hardly do: an organism can't "emit" a change—it emits a response which is in fact a change). The possibility of entering into a jungle like the one in which the fractional anticipatory goal response was last reported begins to appear. One can read the latent learning controversy directly as a battle over the workability of various definitions of "response," including implicitly Skinner's. Kendler's paper (4, pp. 269–77) illustrates the connection well.

This process of redefining in order to save the regularities is a very important, legitimate, and useful part of the procedure of theory-building. The history of the concepts of "temperature," "intelligence," and, more recently, "anxiety" affords some striking examples of this point. Why did we abandon the definition of temperature in terms of the ideal gas scale? Principally because we came to see that the behavior of no physical gas could possibly give us that smooth monotonic curve, which the background theory insisted on for a base line, in the regions to which thermodynamics had turned. It is significant of an inadequate realization of this redefinition maneuver that, in his attack on mentalistic and psychoanalytic explanation, Skinner twice produces what he takes to be terms dating from early animistic explanations, both of which have been transformed as the general theory incorporating

* See Skinner's comments on the great difference between sixty-second and fifteen-second intervals in producing "superstitious" behavior in pigeons (SHB, p. 85).

them has been transformed. Both "force" and "visviva" have been respectively redefined as the product of mass and acceleration and twice the kinetic energy, and the view that their original use involved explicit animism is as speculative as the view that Skinner's use of the term *strength* (of a response) involves the same commitment.* It is not even true that Skinner's terms are, in any sense, operationally definable, as we shall see later.

Skinner starves his pigeons to 80 per cent of their *ad lib* weight; he is fortunate that such a simple procedure yields workable results. We may expect, with some confidence, that with higher organisms this regimen will not yield quite such high intra-species predictability. The pigeons show greatly different response rates but the same type of learning behavior under these conditions; apes may display basically different learning behavior in the same circumstances. Skinner would be ready to introduce new qualifications to the laws (perhaps by widening his "emotional behavior" escape valve) or so to extend the meaning of his terms as to provide the greatest generality. It is instructive to consider the differences between SHB and *The Behavior of Organisms* published fifteen years earlier, in 1938. For an author who, in the earlier book, considered himself to be almost above producing explanations or hypotheses (p. 44), there have been substantial changes in the description of behavior. They are not due solely to early error, but rather to a present preference for a different formulation—for theoretical reasons, ones which were by no means wholly absent in the 1938 discussion (see W. S. Verplanck's chapter on Skinner, (11)). This is the theoretical approach. Hull was supposedly operating in the same way, Hebb and Tolman too, and, as we shall see, Freud. It is comparatively unimportant methodologically that some of them start by trying, or talking about, a more ambitious goal than Skinner's—i.e., a tie-in with neurology; it is a matter of practical interest if, by so doing, they fail to achieve the primary goal. What is important is that Skinner's terms are defined and redefined for purposes of optimizing their value in a systematic account of behavior, that his observations are referred back to a hierarchy of laws for explanation in an orderly way with the Law of Effect forming the keystone, and that his intentions are of a certain kind, namely, the discovery of causal (functional) laws. These

* For a further discussion of this point, see the discussion of "metaphorical devices" below.

are the marks of the theorist. I repeat that I am not here saying anything about the more moderate Skinnerian claim that the neural, mentalistic, and conceptual theorizers have not done as well in discussing learning theory as Skinner has. I am stressing the fact that this is a dispute between proponents of various theories, albeit theories of different levels of abstraction, strictly comparable to that between the proponents of classical and statistical thermodynamics, between macrochemistry and the valence theorists, and to a dozen other disputes in the history of science. I wish to argue that there is no logical or methodological error involved in the approach to behavior through the C.N.S. (whether the "C" stands for "Conceptual" or "Central") or the drive-reduction model, unless we wish to argue that trying to run before we can walk is (a) a relevant criticism and (b) a methodological error. Even there, especially in view of the latent learning experiments, it is difficult to establish conclusively that the time for running is so far away that expectancy theory never got moving.*

Before I conclude this discussion of learning theory, it is perhaps worth commenting on the very casual way that the term "learning" itself is defined by the theorists. Skinner, an environmentalist to the end, says, "We may define learning as a change in probability of response . . ." (ATLN, p. 199). No account of learning which is part of a Skinnerian science of human behavior will ever accomplish this, since the changes in the probability of a response following a disease infection such as chicken pox (increased strength of the scratching response, etc.), an injury, involutional changes, and genetic changes are all forever beyond the reach of a molar theory of psychology, although it is by no means easy to exclude them by a watertight definition. Indeed, from this difficulty springs another justification for a subcutaneous rounding out of the theory of learning—via the attempt to explain variations (intra- or inter-individual) in the capacity to learn or in the basic laws—just as the difficulty of excluding genetics from evolutionary theory makes the attempt arbitrary and costly. As Skinner

* Perhaps I should add the comment, quite irrelevant to the methodological discussion in which this paper consists, that, in my view, the practical achievements of the Skinner-box (and the teaching machine) and the theoretical analyses of their inventor will rank as more lasting contributions to learning theory than his methodological arguments. But if Skinner's methodology is unsound, it is in part a needed reaction against the colossal failure of Hull's theory-building attempts. In his review of "Principles of Behavior" (7) Skinner presents some insuperable criticisms, that only in their later overgeneralization become invalid (ATLN, SHB, CPC'1).

admits, in a revealing passage with which I shall conclude my discussion of his approach to learning,

A biological explanation of reinforcing power is perhaps as far as we can go in saying why an event is reinforcing. Such an explanation is probably of little help in a functional analysis for it does not provide us with any way of identifying a reinforcing stimulus as such before we have tested its reinforcing power upon a given organism (SHB, p. 84).

This implies that appropriate developments in biochemistry and physiology could even make them useful for a functional approach; i.e., Skinner is here conceding that there are no methodological errors involved in micro-theorizing.

Skinner's critique of psychoanalytic theories and concepts, to which I now wish to direct my arguments, is based on certain extensions of the basic position considered above.

One can understand his approach better if one bears in mind two points which are already pretty clear in his discussion of learning theories. First, he is almost exclusively concerned with the applications of a study of behavior, the control aspect, and he is consequently somewhat intolerant of considerations which do not at the moment display signs of giving us more control over behavior. Second, the self-application of the first point makes it very important to him to state his arguments as forcefully as possible—to convince people to change the way they spend their research hours, the way they think. I shall argue that the first of these points makes Skinner too impatient with psychoanalytic theory, and that the second makes him overstate his case, leads him to put it in terms that suggest logical error rather than lack of practical consequences. And, surely, it takes a man strongly affected by both these considerations to say, "When the individual is wholly out of control, it is difficult to find effective therapeutic techniques. Such an individual is called psychotic" (SBH, p. 380). Psychoanalytic terminology has its defects indeed, but it is surely not quite so outcome-oriented.

My specific contentions will be that Skinner, in general, underestimates (a) the empirical content and (b) the practical utility of propositions about 'mental states,' including unconscious ones, and that he overestimates the commitments of psychoanalytic theory and hence the deleterious effects of Freud's influence.

Let us consider some of Skinner's proffered analyses of 'mental-state' ascriptions. At a very general level, he considers descriptions of purpose to be open to the same objections, and I shall begin by taking one such example.

... we ask him what he is doing and he says, "I am looking for my glasses." This is not a further description of his behavior but of the variables of which his behavior is a function; it is equivalent to "I have lost my glasses," "I shall stop what I am doing when I find my glasses," or "When I have done this in the past, I have found my glasses." These translations may seem unnecessarily roundabout, but only because expressions involving goals and purposes are abbreviations (SHB, p. 90).

Without agreeing or disagreeing with the thesis that "expressions involving goals and purposes are abbreviations," we may show that such expressions involve (i.e., are equivalent to, or imply) a great deal more than any one of Skinner's suggested "translations." If a man correctly describes himself or another as "looking for his glasses," then we can infer *all* of the following:

- (1) He owns glasses or believes he does.
- (2) He does not now know where they are.
- (3) He is engaging in operant behavior of a type that has previously led him or others of whom he knows to find objects of this kind (Skinner's restriction to the man's own experience with lost glasses is clearly too narrow).
- (4) The aspect of his behavior described under (3) will cease when he discovers his glasses.

There are certain 'mental-state' references in (1), (2), and (3). These, in turn, may be reduced to Skinnerian terminology (with or without loss—I am not now passing judgment on that aspect of the issue). To say that the man believes he owns glasses is necessary, since one could not correctly describe someone as "looking for his glasses" unless he believed that he possessed glasses. It is too strong a condition to require that he actually own glasses since they may, in fact, have fallen into the wastebasket and been long since consumed. But we do not have to imagine that beliefs are mysterious states of the man's forever unobservable mind; we can, in turn, reduce "X believes Y" to a series of statements, at least some of which are conditional and all of which involve probability of specified observable responses. We can, in fact, avoid the conditional element (again, perhaps, at some cost) by re-

stricting the type of prediction we make. Without going to this extreme (the arguments for so doing are entirely philosophical, since the conditional has an extremely respectable place in the logical vocabulary of science and mathematics), we can expect the analysis of "X believes Y" to involve a series of statements such as the following.

(a) If X is asked "Do you believe Y?" under standard conditions, he will reply "Yes" or "Indeed" or "Of course" with 80 per cent probability.

(b) If X is hypnotized or injected with specified amounts of sodium ethyl thiobarbiturate, he will respond as in (a) with 75 per cent probability.

In the same way, we can analyze statements (2) and (3), in which the word "know" occurs.

Two comments on these subsidiary analyses should be made. In the first place, each of the samples given contains arbitrary probability figures, presumably not based on research. How could one give a more precise analysis? Secondly, giving a complete list of the stimulus-response patterns that would be relevant is clearly a problem of equal difficulty. It is in the face of the peculiar intransigence of these problems that Skinner, amongst others, is led to abandon the entire vocabulary of 'mental-state' descriptions. Why not, he says, stick to the unambiguous observation language in view of the remarkable achievements possible within its boundaries (the moderate position) and the eventual necessity of returning to it, assuming that one could successfully complete the analysis of "believe," "know," etc. (the extreme position). The reply is that, if we are seeking to analyze any purpose, goal, desire, intention, or mentalistic language, we can succeed only by doing the analysis, not by considering another problem; it should be clear from the discussion above that Skinner's suggested "translations" are thoroughly inadequate.

This is not to deny that Skinner can do a great deal, much of it related to this problem, without actually solving the problem. One might ask whether the original problem is really worth troubling with, when we come to see the extreme vagueness of ordinary language. But the original problem and its relatives make up the problem of analyzing purposive behavior, i.e., what we ordinarily call purposive behavior (or mentalistic language of the other varieties mentioned); and we can't palm off an analysis of something that our intuition suggests is

the "essential behavioral component" instead, because intuition isn't objective argument and people have widely varying intuitions on this matter. In fact, the subtleties involved in ordinary language are considerable and, as I shall later argue, highly functional. The arguments so far given are intended to support the thesis that Skinner's analysis is much too simple, and that a complete reduction to Skinnerian language ('operationally' definable terms in response probability statements) is extremely difficult. Now, before we conclude that some simplified reduction will have to be made if any scientific investigation is ever to be possible, let us examine some of Skinner's own "operationally definable" language and see whether it is really free of what we can call (after Waismann) the "open texture" of the ordinary language of intending, knowing, and believing, and a fortiori of unconscious attitudes. If it is not, we shall have to decide whether the gain involved in the "simplification" is actually sufficient to compensate for the confusion involved in changing the meaning of many familiar terms. After all, we could not get very far by defining "intelligence" as "the average rating on a 1-10 scale scored by three friends"; the proposed operational definition corresponds very poorly to our background notion of intelligence since it will not be at all constant for each individual—and it doesn't analyze the notion itself, merely referring it to others for analysis. Furthermore, it is spuriously operational since there is no one such rating, and it is a spurious improvement for all these reasons. Are Skinner's "reductions" substantially better than the original terms? He makes a poor start with "intelligent," of which he says it "appears to describe properties of behavior but in reality refers to its controlling relations" (SHB, p. 36). The reality is that it describes a property of behavior, which does not at all prevent its analysis from involving reference to its controlling relations. Skinner's belief in the incompatibility of these alternatives presages an unsound analysis of the permissible and useful modes of definition.

Consider another example: in CPCT Skinner refers to

... another pitfall into which Freud, in common with his contemporaries, fell. There are many words in the layman's vocabulary that suggest the activity of an organism yet are not descriptive of behavior in the narrower sense. Freud used many of these freely—for example, the individual is said to discriminate, remember, infer, repress, decide, and so on. Such terms do not refer to specific acts. We say that a man discrimi-

nates between two objects when he behaves differently with respect to them; but discriminating is not itself behavior. We say that he represses behavior which has been punished when he engages in other behavior just because it displaces the punished behavior; but repressing is not action. We say that he decides upon a course of conduct either when he enters upon one course to the exclusion of another, or when he alters some of the variables affecting his own behavior in order to bring this about; but there is no other "act of deciding" (p. 304).

Looking at this example of Skinnerian analysis, we can again ask (a) whether or how nearly the products of the analysis are equivalent to the original, and (b) whether they are really acceptable for Skinner's purpose, i.e., operationally pure. In this instance, I wish to concentrate attention on the second question. Skinner's analytic base here includes the following crucial terms: "behaves"; "behaves differently with respect to X and Y"; "behavior segment X displaces behavior segment Y"; "enters upon X to the exclusion of Y"; "alters variables affecting own behavior." Elsewhere in this paper, we find him using (not quoting from Freud) such terms as "reinforces," "verbally injures," "pattern of behavior," etc. Now, we all know pretty well what these terms mean. But it must be noted that to say an organism "behaves" is not to describe its behavior in Skinner's "narrower sense"; and to say that some of its behavior "displaces" other behavior is no more to describe behavior in Skinner's sense than to say that one course of action was decided on rather than another (unless one makes the erroneous assumption that the latter necessarily refers to some inaccessible inner activity, which Skinner would scarcely wish to do). It appears that one could reply to this complaint by giving a list of actions or activities of an organism such that the disjunction of them all is equivalent to "behaves." But, clearly, one could never complete such a list. Alternatively, then, could one not equate "behaves" to some concatenation of anatomically-physiologically defined movements? But this is a very awkward alternative, since there are clearly certain movements, like the patellar reflex, PGR, epileptic frenzies, parietic and aphasic behavior, which cannot be explained at all within molar psychology; and there are many 'motionless' states, ranging from the catatonic to the reflective, which Skinner thinks are susceptible to his type of explanation. It is significant that Skinner never attempts to define "behavior" in his book *Science and Human Behavior*. After all, the anatomical-physiological definition would be

likely to create great difficulties in excluding physiological explanatory theories; and the alternative of a long list definition would, equally clearly, be interminable, especially since the progress of a science of behavior will itself bring about the extension of the list. Even more serious difficulties arise over the term "displaces"; deeply embedded as it is in a spatial analogy, it has in addition all the problems involved in defining its range of application, viz. segments of behavior.

Skinner constantly reiterates his complaints about certain metaphors and their awful effects ("When one uses terms which describe an activity, one feels it necessary to invent an actor . . ." [CPCT, p. 304]; qualified by "The point is not that metaphor or construct is objectionable but that particular metaphors and constructs have caused trouble and are continuing to do so" [CPCT, p. 301—the rest of the paper makes it fairly clear that he doubts the utility of any metaphors or constructs in a science of behavior]). It is as well to remember that the whole language of reinforcement, displacement, satiation, etc. used by Skinner is loaded with metaphorical meaning, and meaning to which serious objections can be raised (especially if one takes the very tough line that Skinner does about metaphors—for example, he objects to "motivational interaction" on the grounds that it implies "arrangements or relationships among things, but what are the things so related or arranged?" [CPCT, p. 305]). I have discussed earlier in this paper the utility and legitimacy of using such concepts on a trial basis; their success and that of their subsequent partial redefinition is a matter for scientific appraisal. I want to stress the fact that, although Skinner raises the ghost of his extreme position in CPCT ("It would be difficult to prove or disprove . . . that metaphorical devices are inevitable in the early stages of any science" [p. 301] certainly suggests their superfluity in later stages), he could not possibly avoid these charges against his own system. For amongst "metaphorical devices" he includes the use of terms such as "force" and "essence," on which he says we "look with amusement [as part of the] science of yesterday" (CPCT, p. 301; he elsewhere, as previously mentioned, gives "vis viva" as an example [SHB, p. 27]). It is clear that these terms, which came into the scientific language with some metaphorical connotations, are now entirely respectable terms in mechanics, olco-chemistry, and dynamics ("vis viva" = twice the kinetic energy) and retain the greater part of their original connotations. For example, Skinner says, "The motion of a rolling stone was once

attributed to its vis viva" (SHB, p. 27); it still is. As far as I know, the ancients never imagined the stone was actually alive; they merely attached a name to the hypothesized property which would explain the motion—the property which turned out to be a multiple of the kinetic energy. They chose a word which would carry some connotation of "explanatory of motion" from the field which they best understood, where motion distinguished living things. They were saying, in effect, "Moving inanimate objects have some of the properties of living things; let us name the concept that explains this aspect of their behavior their type of 'life-force.'" It was exactly this line of thought which led to the introduction of the term "response"; there may have been, or still be, people so stupid as to feel it "necessary to invent an actor" who gives the response, but it seems unlikely.

Now, it may be the case that so many people have been misled by the use of terms such as "repress," "sublimate," "project," that the contribution of psychoanalytic theory has been neutralized; but at least the introduction of these terms was as legitimate as that of Skinner's terms. Certainly, at the level of criticism which Skinner introduces, he is no better off; for example, he says, "The notion of a conscious or unconscious 'force' may be a useful metaphor, but if this is analogous to force in physics, what is the analogous mass that is analogously accelerated?" (CPCT, p. 305). Analogies are defined as incomplete parallels: the "force" analogy comes not from mechanics, where the term has been partially redefined, but from the same place that mechanics got it, ordinary usage—and force is not defined as the product of mass and acceleration in ordinary usage but as "strength or energy; vigor" (Webster). Now, if no fundamental logical error is involved in using "metaphorical devices," as I have tried to show, then we should consider the possibility of redefining or reconstructing the theory rather than rejecting it. And, in fact, both Skinner in SHB and Ellis in this volume make serious attempts to do this.

Skinner's objections on this point—the use of "metaphorical devices"—are now, I think, boiled down to the claim that the ones he uses are less misleading than Freud's. And this ties in with his complaints about "Freud's explanatory scheme," to which I shall return in a moment. I hope to suggest that Skinner's explanatory scheme applied to psychoanalysis often provides an interpretation rather than an alternative.

The discussion of metaphor was introduced in the course of analyzing Skinner's own terminology for conformity to Skinner's standards. I hope that I have shown its success in meeting these standards is different at most only in degree from that of the terms to which Skinner takes exception; his own terms still retain the property of open texture; i.e., there is the same type of difficulty about categorically defining "behavior" or "displacement" as there is about "purpose" or "belief."

Yet, there hangs above us the pall of smoke from the battle over introspection.* "How can one deny that 'purpose' and 'belief' are words with an inner reference, and sometimes with no external manifestations?" say the introspectionists, while the radical behaviorists fidget at this metaphysics ("the traditional fiction of a mental life" [CPCT, p. 302]). Here I wish only to argue that Skinner's analysis is unsound, even for a behaviorist, and does not achieve what he thinks it will achieve. It is not a special feature of *mentalist* concepts that they cannot be given an explicit unambiguous definition in basic observation language ("the left hand was raised three inches, the head turned to the right about 45°, eye fixation remained constant, etc."); but it is a feature of *all useful scientific concepts*, including Skinner's own. This point is fundamental; and it is difficult to accept because the whole trend of thought since scientists really became self-conscious about their definitions, say with Mach, has been in the opposite direction. Indeed, it sounds reasonable and admissible to insist on terms that can be explicitly and unambiguously defined in terms of basic observations. It is, in fact, a valuable exercise to attempt this at any stage in the development of a science; but, if successful, one is merely taking a still photograph of a changing scene, and the motion, not the snapshot, shows progress. Both the philosopher and the scientist can learn from the snapshots, but they will not understand the changes without a great many snapshots and a good deal of inference. So it is with the changing meaning of scientific concepts: at any stage it is possible to give their cash value in terms of observations, but to understand them properly

* Skinner's logical maneuvers in his skirmishes with introspection repay careful study. His case depends on a crucial misuse of perception language, e.g., a systematically ambiguous use of "see" (SHB, pp. 273-78). It is further weakened in CPCT by his argument that Freud's evidence for self-deception in introspection shows introspection not to be like observation. On the contrary, it shows how very much it is like observation (itself fallible), though it is not the observation of logically inaccessible inner states, but of states some of whose aspects are sometimes not externally displayed.

one must also know their role in the theory and have some idea of their future movements given certain contingencies. A term is fruitful only if it encourages changes in its own meaning; and, to some considerable extent, this is incompatible with operational definition. It is sometimes easier to learn how to use terms than to learn how to work out their cash value in a certain currency, and sometimes the reverse. We all know very well how to use terms such as "purpose" and "belief" and we can teach someone who doesn't speak the language how to use them (child or foreigner), but it's not easy to reduce them to Skinner's currency (and, some suspect, not possible). But Skinner won't pay in other currency, which he views with suspicion; and, if we want to collect the real value of our investment from him, we must compromise: collect a good deal from him on his terms and talk him around on the rest, perhaps by showing him some defects in his own coinage.

On Skinner's own terms, his analyses are not satisfactory; and, when we look at the currency itself, it seems to have the same weaknesses as the promissory notes that he won't negotiate.

There were two difficulties about our original analysis of "belief," which sprang from an attempt to improve Skinner's analysis of "looking for his glasses." The first, into which we have now gone at some length, stems from the impossibility of completely itemizing the units of behavior to which "belief" (in some way) refers, and is a difficulty which some of Skinner's basic terms share. The second was the difficulty of giving precise estimates of the response-probability in any of the component statements. He would, I think, be the first to agree that the corresponding statements in his own system—i.e., dispositional statements involving response induction or stimulus discrimination and generalization—could not contain exact values except by arbitrary redefinition. Corresponding to "S believes p" would be "S discriminates red keys from green keys in a Skinner-box." We had difficulty in giving a cutting frequency for affirmative responses (to the question "Do you believe p?" put to S as a stimulus) which would differentiate a state of belief in p from one of uncertainty about p. Similarly, Skinner could not guarantee in advance the response-frequency or, indeed, the existence of discrimination, if we set up conditions different from those under which the S was trained—for example, there is the possibility of what he terms "emotional behavior."

A great deal more could be said about the comparison here with

reference to the Behaviorist's philosophical thesis, but I wish to avoid direct engagements with this and instead bring the above points to bear on Skinner's account of Freud's "explanatory scheme." I would conclude this examination of Skinner's attack on purposive and mentalistic terminology by saying that the substitute he offers is not obviously required, is not as good as it could be, and, even if it were improved to its natural limits, would be insufficiently distinguishable from the original to justify the effort. Skinner is practically allergic to even the most harmless references to the 'mental life'; and I think this prevents him from seeing that, if his translations were really satisfactory, then there would be a 'mind' in his system, too. Consider, for example, this quotation: "One who readily engages in a given activity is not showing an interest, he is showing the effect of reinforcement" (SHB, p. 72). But a very special type of effect, one which it is extremely important to distinguish from the readiness to engage in a given activity springing from severe punishment for failure to do so. One can no more deny that Skinner has shown an interest in psychology during much of his academic life—i.e., the propriety of the phrase "showing an interest" in this case—than one can deny that this book has pages (i.e., since standard conditions obtain, the propriety of the phrase "books have pages") even if it does suggest ownership. We shall see how this hypersensitivity to one interpretation (quite possibly an unfortunate or unproductive one) leads him to underestimate the utility of psychoanalytic theory. It is by no means necessary to interpret these phrases in this way; and, in fact, we can improve Skinner's suggested behavioral analysis to the point where it provides a good scientific substitute for the original phrase. This could be described as 'translating the mind into behavioral terms'; or, equally well, it could be said to show that the reference to mind was not objectionable in the first place. To say that such phrases as "making up one's mind," "having an idea," etc. are "obvious" cases in which the mind and ideas are "being invented on the spot to provide spurious explanations" (SHB, p. 30) is to miss the point that these are genuine explanations but do not involve reference to some scientifically inaccessible realm. No one, without stretching an etymological point, reads into such phrases what Skinner objects to in them; they are invariably used as shorthand for a set of descriptive plus dispositional propositions, just as "believe," "behave," and "beta-particle" are.

Taking the bull by the horns, I shall argue that Skinner, partly be-

cause of errors of translation of the kind just discussed, underestimates the practical value and logical validity of explanations of behavior in terms of 'mental events.' Basic to all his arguments is the belief that any reference to inner states is an unnecessary complication, since eventually one has to explain *them* in terms of environmental variables. The counterargument that the egg comes before the chicken and that we must, therefore, get back inside the skin to the organism's inherited genetic composition in order to achieve a complete explanation would presumably cut no ice because, Skinner would say, we can't control the genes (or neurons). But we mustn't be misled by the combination of Skinner's use of the word "control" and Skinner's analyses of familiar processes. For example, he argues that "awareness of cause [of one's actions] has nothing to do with causal effectiveness" (CPCT, p. 305). But this is a verbal trick. He is not really denying (as he appears to) that the patient who attains insight into his aggressive remarks to his brother will often abandon them subsequently. He is (essentially) arguing that, since both insight and behavior improvement are due to changes in other environmental variables, it is wrong to say one of them causes the other, i.e., wrong to say that it is the awareness of the hostile impulses that reduces their effectiveness. He is not denying that the first might always be followed by the second or that the second might never occur without the first, i.e., that the process of bringing about the insight produces the remission of the symptoms. Now, analysts, including Freud, have long stressed the fact that mere enunciation of the relation between unconscious hostility and aggressive remarks will not bring an end to the latter. Thus the analyst believes

(1) The hostile feelings (C_1) cause the aggressive behavior (E_1).

(2) Getting the patient to the stage where he volunteers the interpretation—i.e., getting-him-to-achieve-insight-in-this-respect (C_2)—causes the symptoms' disappearance (E_2).

Now, achieving insight is the standard psychoanalytic case of "awareness of cause," and it is followed by the termination of the causal relation (1); moreover, the process of achieving insight actually causes the improvement (as stated in (2)). It is, therefore, extremely misleading of Skinner to say "awareness of cause has nothing to do with causal effectiveness." It is trivially true that if A causes B under conditions C, and C does not contain a provision excluding an observer, then A will still cause B even if there is an observer. The analytic discovery

was that, in certain cases of caused behavior, C does include such a prohibition and A no longer produces B in a patient P when P has learned about the connection (this schema oversimplifies in certain respects). It might appear that Skinner is making a more substantial point, that he is suggesting something empirically distinguishable from Freud. I am not clear that he is. The following quotation suggests there is no such tangible difference, but only a difference of emphasis: "Therapy consists, not in getting the patient to discover the solution to his problem, but in changing him in such a way that he is able to discover it" (SHB, p. 382). But on other occasions this is not clear: "The parallel between the excision of a tumor, for example, and the release of a repressed wish from the unconscious is quite compelling and must have affected Freud's thinking" (CPCT, p. 301)—and Skinner has made clear to me in conversation that he regards this as perhaps Freud's most serious error. Now, what is the pattern of explanation here? We have

1. The tumor (C_1) causes illness (E_1)
2. Removal-of-the-tumor (C_2) causes recovery (E_2)

and the analogy, which Skinner disputes, would be

- 1'. The repressed wish (C'_1) causes, say, stuttering (E'_1)
- 2'. Release-of-the-wish (C'_2) causes cure (E'_2).

If Freud thought that the "releasing of a suppressed wish from the unconscious" meant the mere uttering of the words, perhaps even parrot-wise, after the analyst, and was in itself curative, then Skinner is disagreeing with Freud on empirical grounds. It seems clear enough that this was not Freud's view. But if Freud thought that the mere verbalization had little or no therapeutic value, whereas the 'spontaneous' release of a wish marked the penultimate stage of therapy for this symptom (culminated by the application of this insight), then I am not sure that Skinner would disagree, in view of the fact that he is presenting no new evidence. After all, discovering the wish makes it possible for P and the analyst to reorient P's behavior accordingly, and the fact that it has been voluntarily produced makes it likely the behavior will be successfully reoriented. If the reorientation is successful, could one not then say that producing the wish produced the cure under these circumstances? The hard work goes into achieving C'_2 , but if we define cure as the actual vanishing of symptoms, then it is the release of the wish

that produces it. Thus, we might agree with Skinner that therapy consists in changing the patient until he can achieve C'_2 , but that the (proximate) cause of the cure is C'_2 . And Freud would not, as far as I can see, disagree with this. I do not think this position justifies such strong statements as the following:

Freud's contribution . . . [was] . . . not that the individual was often unable to describe important aspects of his own behavior, or identify important causal relationships but that his ability to describe them was *irrelevant* to the occurrence of the behavior or the effectiveness of the causes (CPCT, p. 304; my italics).

I have gone into this point in detail because I think it well illustrates the difficulty of dealing with Skinner's criticisms of Freud. An attempt to wholly disentangle his ideas about the formal status of psychoanalytic theory would be unrewarding, but I think it important to show that it rests on an erroneous dichotomy, which explains a great deal about the rest of his approach. He says, "No matter what logicians may eventually make of this mental apparatus, there is little doubt that Freud accepted it as real *rather than* as a scientific construct or theory" (CPCT, p. 301; my italics).

Skinner is here reaping the whirlwind of early positivism, and this eddy affects his position more seriously than his other philosophical inheritances from the same source.* The idea that scientific constructs are not "real" but are mere "explanatory fictions," as he goes on to describe them, is untenable; but, if one believes it, one will not be encouraged to invent many. It is a little misleading to insist that scientific constructs are real, although it is certainly better than the alternative (I have never heard anyone arguing whether theories are real: what would an unreal theory be like?). The issue is a spurious one; such constructs should always be said to be real in *such-and-such* a respect, but *unreal*—i.e., unlike such earthly real things as platypuses—in *such-and-such another* respect, etc., etc. Few are observable like animals but many have observable consequences, like the neutrino, and their existence is said to be confirmed when these consequences eventuate. Philosophy has outgrown such questions as "Do groups exist over and above their members?" "Do electrons really exist?" though it learned much

* E.g., "Certain basic assumptions, essential to any scientific activity, are sometimes called theories. That nature is orderly rather than capricious is an example" (ATLN, p. 193).

from answering them. The answer must always be, "In a certain sense, yes, and in a certain sense, no." The important point is that a theory can and does proceed without having to distinguish each construct as "real" or "explanatory fiction"; there are other rules to be obeyed (those of factor-analysis in certain cases, for example) but not this rule. We may have grave reservations about the meaning, if any, of "libido," which we may awkwardly express by asking whether there really is a libido; but if a satisfactory answer can be given, it does not show that "libido" is not a scientific concept, that it is an observable: it would be a scientific construct, it would 'exist,' but it need not be actually observable as long as it has effects we can observe. The libido has to undergo the same examinations as Spearman's *g*, the concept of isostasy, and cultural lag. If they are satisfactory constructs, they may figure in explanations in a way we shall consider, although they are not observables and not all operationally definable in terms of observables. But that does not make them "fictions"; indeed, "explanatory fiction" is a contradiction in terms: abbreviatory devices are not normally explanatory, and explanatory devices are never fictitious.

There is no obligation on the libido to be located in the brain, or to have spatial location at all; the rules are flexible, they require only that we be able to make some objective distinction between cases in which the term can be properly applied and the rest. The same applies to Skinner's use of "reserve" (of responses—ATLN, p. 203) or the physicist's concept of entropy. But Skinner believes in the "observable-or-explanatory-fiction dichotomy," so he is at great pains in SHB to show that introspection is not a form of observation and, hence, that "conscious, as well as unconscious events [are] inferences from the facts" (CPCT, p. 302)—hence, are constructs and hence *unreal*. On one occasion, he puts it in this way: "... the act of self-observation [sic] can be represented within the framework of physical science. This involves questioning the reality of sensations, ideas, feelings, and other states of consciousness . . ." (CPCT, p. 304). The example which follows shows, as one might expect, that he is not going to question the *reality* of these things but the *reliability* of reports on them: the mystic's religious experience is not *unreal* but (according to Skinner's hypothesized psychoanalyst) incorrectly described as communion with a supernatural being. One can often show that someone's description of his sensations is inaccurate; it does not follow that he had none.

Having arrived at the conclusion that mental events are *unreal*, Skinner feeds it back into the argument and attacks explanations in terms of mental states as circular and/or superfluous. Only in this way, as he sees it, can one get past the Scylla of explaining how the mind and body interact and the Charybdis of explaining how we can know the contents of another's mind. Union rules prevent me from divulging the secret of the problems which provide a guaranteed annual wage for philosophers; but it is not too difficult to show some very important ways in which mental states can figure in extremely respectable explanations, in whose company, indeed, no behaviorist should feel ashamed to be seen.

Suppose that we accept the behaviorist analysis of mental states that Skinner is anxious to sell us, making some modifications to the actual model he is vending, along the lines suggested in our account of "He's looking for his glasses" and "He's showing an interest in this activity." Ignore, for the moment, the twin difficulties of inexhaustibility and inaccuracy in specifying the exact propositions in behavior-language to which we reduce the mental state. Then, taking a simple case,

"X is in mental-state M." = "If X is in circumstances Y, he will (with probability P) do Z."

Now, there will normally be many other circumstances (Y' , Y'' , . . .) under which X will also do Z. For example, he will drink a glass of water (Z) not only when thirsty (M)—i.e., when he has been deprived of water for some time, etc. (Y)—but also when there is a gun at his back even if he is satiated (Y'), when he has been deprived of food and no food is now present (Y''), etc. Suppose that X is in an observation room, taking part in an experiment of whose purposes you are ignorant. You do not know whether he is being paid, deprived, or intellectually stimulated. At 3:00 P.M. he looks at the clock, takes up a glass of water, and drinks it. You ask, having these other possibilities in mind: Why did he drink? A perfectly proper and informative explanation is "Because he is thirsty." It is informative because you now know (always assuming the behaviorist analysis to be correct) that this means circumstances Y must obtain, as opposed to Y' , Y'' , . . .

This type of request for an explanation is perhaps the most common of all. Skinner would describe it as a request for information about "the external variables of which behavior is a function" because he thinks

that 'inner-state' language is really translatable into descriptions of external variables. But if it is, then explanations in terms of inner states are perfectly legitimate. And even if it is not, even if there is something more involved in reports on inner states, it is still true that they are connected with external variables, i.e., whatever "X is thirsty" may refer to, we know it is probabilistically connected with hours of deprivation, aversive preconditioning, etc., so that the worst charge that Skinner can lay is vagueness. Of course, he may not like the obscurity of the connection between a state of thirst and behavior; we can help him by accepting tentatively the identification of the two. Thus, one does not have to solve the mind-body problem in order to find out what independent variables are related to subjective reports of feelings of thirst, actual drinking without such reports, etc.; and even the decision that thirst is not *wholly* reducible to behavioral criteria does not imply that it is not a good basis for explanation, since it clearly has many behavioral consequences. It is not important to a public health officer in charge of mosquito control whether we define DDT in terms of its entomological effects or whether we define it chemically, as long as we don't deny its entomological effects; similarly, whether we say that thirst is or is not something more than a certain pattern of behavior is unimportant, as long as we agree that it is the psychologist who studies the *whole behavioral aspect*. The presence of vagueness not only affects his own more general concepts (notice the difficulties in distinguishing operant from response conditioning under certain circumstances) but also has certain virtues for a scientific concept in the field of behavior, a fact which will be further elaborated below.

The first point made, then, is that, on a radical behaviorist analysis of inner states, explanations in such terms are vital and legitimate. I did not say "on Skinner's analysis" because we have already made a number of improvements on that without abandoning radical behaviorism. To uncover some more of the very serious difficulties in Skinner's own analysis, I am going to examine one or two further instances of it, still on his own standards. Then I shall consider the results of our suggested changes in the radical behaviorist analysis of mental-state terminology, in an attempt to show that one can produce a fully scientific account that is much closer to being an analysis of the actual mental concepts we ordinarily employ, which Skinner views with such suspicion.

But what is Skinner's position?

To what extent is it helpful to be told "He drinks because he is thirsty"? If to be thirsty means nothing more than to have a tendency to drink, this is mere redundancy. If it means that he drinks because of a state of thirst, an inner causal event is invoked. If this state is purely inferential—if no dimensions are assigned to it which would make direct observation possible—it cannot serve as an explanation (SHB, p. 33).

All these conclusions are erroneous. Even if "to be thirsty means nothing more than to have a tendency to drink," it is by no means merely redundant to be told that *on this occasion* he drank because of that tendency *rather than* under compulsion or because of a tendency to eat or etc. To have a tendency is to have a certain disposition, and everyone has some disposition to drink, but it is not always that disposition which explains our drinking. When we are satiated, for example, we have a short-term disposition not to drink; and, in such a case, the statement that we drank because we were thirsty would not, even on Skinner's first analysis, be redundant; it would be false. It follows that Skinner does not see the importance of dispositions, nor does he see the nature of what I shall call "discrimination-explanations," i.e., explanations of event E as due to antecedents A *rather than* A' or A'', all of which we realize are capable of producing E.

Of course, there are occasions on which pseudo explanations—looking rather similar to this one—are offered. If we explain someone's frequent sleepy appearance by saying he is a soporific type, we may well be deceiving ourselves. But to chastise the ordinary explanations of *individual* events by reference to dispositions, on the grounds that some explanations of *patterns of behavior* by reference to dispositions are redundant, is manifestly unfair. Skinner's failure to distinguish these is clearly shown in the following quotation:

When we say that a man eats *because* he is hungry, smokes a great deal *because* he has the tobacco habit, fights *because* of the instinct of pugnacity, behaves brilliantly *because* of his intelligence, or plays the piano well *because* of his musical ability, we seem to be referring to causes. But on analysis these phrases prove to be merely redundant descriptions. A single set of facts is described by the two statements: "He eats" and "He is hungry." A single set of facts is described by the two statements: "He smokes a great deal" and "He has the smoking habit." . . . (SHB, p. 31).

This is guilt by association! Every one of these examples except the

first is susceptible to Skinner's criticism; but the first is the vital one for his attack on mental states, partly because we don't introspect "having a smoking habit" whereas we do introspect "being hungry." It is a logical error, even for a radical behaviorist, to imagine that the same "set of facts" is described by "He eats" and "He is hungry" since, for a radical behaviorist, the second statement is equivalent to "He has a disposition to eat," i.e., "Under specifiable conditions it is P per cent probable that he will eat," which neither implies nor is implied by "He eats."

Skinner's third conclusion—that if thirst is a purely inferential state not susceptible to direct observation, it cannot serve as an explanation—is also in my view mistaken. I have argued above that it is only necessary for the hypotheses embodying scientific concepts to be susceptible to confirmation, not observation—or else temperature, inertial mass, and the spin of the electron would be illicit—and the argument applies here. Furthermore, it is possible to construct an example which will demonstrate this and, at the same time, answer any doubts that may have arisen in the reader's mind as to the propriety of explaining one observed drinking-event in terms of a disposition if the latter is interpreted as a construct out of drinking-events.

Suppose that we introduce the symbol Ω into psychological discourse in the following way: if an organism O is such that variable $v_1 > N_1$, while $v_2, v_3, \dots, v_n < N_2$, we shall say that O is in state $\Omega(\Omega(O))$. Further, if $v_1 \leq N_1$ while $v_2 + v_3$ or $v_2 + v_4$ or $v_3 + v_4$ or $\dots > 2N_1$ and $v_1, v_2, v_3, \dots, v_n < N_2$, we shall also say that $\Omega(O)$. And if some one of $v_2, v_3, \dots, v_n \geq N_2$, we shall say that it is not the case that $\Omega(O)$. Notice that we have not said anything about the case $v_1, v_2, v_3, \dots, v_n \leq N_1$, except under certain conditions on the ratio N_1/N_2 between the constants. This example is a rough model of a personality or pathology category, interpreted operationally. Now, we have not in any sense assigned dimensions to Ω , any more than we do to schizophrenia. According to Skinner, then, "it cannot serve as an explanation" since no direct observation is possible. But it may provide a most useful means of classifying organisms, such that laws can be stated in simpler forms—i.e., it can be a step toward a sort of theory which Skinner approves of ("a formal representation of the data reduced to a minimal number of terms" [ATLN, p. 216]), and it can certainly figure in discrimination explanations. What can Skinner say to resolve this con-

tradiction? The necessary compromise seems obvious: Ω is satisfactory because it at least is defined in terms of observables. So it appears. And yet the word "state" occurs. Can we really infer from any formal considerations that O is in a certain state? Indeed not: one could not introduce a state defined by reference to the number of people within a one-mile circle with center fifty miles south of O without changing the present meaning of "state." "State" is a word that needs a great deal of unpacking; it is appropriate only when we are sure that a causal account of the phenomena that differentiate "states" can be given in terms of actual physical changes in the organism corresponding to changes of "state" in the proposed sense.

This simple fact makes the language of states at once more complex and more useful than Skinner allows, and it partially explains how the tie-in of psychoanalysis or psychology with neurology is so strong. As it stands, Ω can give us discrimination-explanations of the fact that certain variables have certain values (e.g., that X drinks or stutters or produces neologisms) insofar as it contains a reference to other variables which are causally related to those observed, and it can only do that by referring to a current state. Even the thirst case, which appears as the simplest possible example, where v_1 is the only variable on which Ω depends as well as being that which is explained, is less simple than it appears since, for the explanation to be valid, it must be the case that the earlier history of v_1 variation has causally effected the present state of O. It is immaterial that we cannot demonstrate this physiologically. As long as we have no direct evidence against it and it works as an explanation, then we have evidence for it, indirect evidence. But it must be true. Moreover, it is not incumbent on us to say how the matter could be tested by direct observation at all. We may believe that this will be possible, e.g., by identification with some neural and intestinal configuration; but we may also believe it will never be possible, as long as we are prepared to give (a) the conditions which will count as confirming evidence for and against the use of the construct, and (b) reasons for believing that some causal connection is possible between the independent environmental variables and the behavioral output variables. Thus, one could not, in the world as it now is, argue that an Ω defined by reference to the past population density of an area geographically distant from O could explain either the present drinking behavior or indeed almost any other behavior

of O, since one can give no reasons for thinking that it could produce any present effects on O at all.

Now, it is this basic process in science—the process of ascribing states to substances and organisms—which forms the first level of theory-building; and it is one that Skinner cannot avoid himself, for example, when he describes an organism as “satiated.” This term would have to be abandoned even if Skinner’s data was still accepted, if we did not believe that the gratification had produced an effect on the organism that, in fact, persisted. Our reason for thinking this is true is the change in response frequency after unrestricted eating (for example) is allowed, when the reinforcement is food. This striking result, we would argue, shows clearly that an effect has persisted on the organism. Let us call this effect “satiation”; it is a state of the organism. Now, Skinner overlooks the theoretical element in this analysis and imagines he can define the word without any reference to state: he views it as a summary of past history. I have already shown that he gives a bad summary; now I am arguing that he is committed to more than a summary. The reason is that one cannot believe the past history to affect the present behavior except via a present state—and science as a whole (not molar psychology) must explain how this implicit hypothesis (supported by every success of molar psychology) is justified. It is a rare explanation which does not produce new discoveries—and one need give no further justification for neurologically-oriented behavior research. Skinner is misled by the argument that neural states are dispensable and inaccessible (for molar psychology) into behaving as though they are scientifically dispensable, which they are not, even if inaccessible, and into imagining that he does not use them. So, states may be “purely inferential” and yet “serve as an explanation.” It is psychologically unfortunate if other people are really misled by talking about states—conscious or unconscious—into imagining that science should not be concerned with discovering the antecedents, but the sin is no worse than that of imagining states to be dispensable.

Skinner is seriously in error on this whole issue, and understandably so—for if he too often acknowledges this implicit and necessary belief in the state-differences brought about by various schedules, he would have to face the question, What sort of state-differences do you have in mind? The reply that he means no more by the state than “that

which is produced by the past reinforcement schedule and which produces the future responses” would, of course, be open to his own objection that it is an “explanatory fiction,” unless he can give directly observable properties to it. Certainly it is obvious that if the organism reacts differently, it is in a different state; but insofar as its reactions are physically determined by its neurology, so far it is obvious that neurology must provide the foundations for this basic assumption of molar psychology. Now it becomes clear why Skinner prefers to talk of functional dependency rather than causal dependency: because causal dependency is necessarily mediated by a state of the organism about which Skinner can say nothing. We cannot object to what Skinner does positively; but in his criticisms he implies that the supplementary activities are unnecessary and invalid, and in this he is surely wrong. Of repression and the other defense mechanisms, he says in CPCT that they should not be regarded “as activities of the individual or any subdivision thereof . . . but simply as ways of representing relationships among responses and controlling variables.” But they are activities (successions of states) of the individual, the very ones necessary to link the controlling variables to the responses; and we can still agree that they are to be distinguished by study of the functional relationships they mediate. Skinner does not hold the only alternative position—that the childhood trauma itself directly causes (across space and time) the neurotic behavior—but at times he sounds very like it. The state of anxiety, he says, “is of no functional significance, either in a theoretical analysis or in the practical control of behavior” (SHB, p. 181). Well, it is an absolute necessity in the total scientific study of behavior to have such states, and the anxiety-reducing drugs show that this one has considerable functional significance in the control of behavior, somewhat contrary to the implications of this passage or of the even more dogmatic statement preceding it—“Any therapeutic attempt to reduce the ‘effects of anxiety’ must operate upon these [controlling] circumstances, not upon any intervening state.”

I want now to suggest that the psychoanalytic or group dynamic or historical or sociological approach is an alternative (to molar psychology) of a most respectable kind for dealing with certain areas of behavior. From a certain point of view, these approaches are simply examples of molar approaches: they are molar approaches to molar behavior in certain areas—they are branches of molar behavior science. The psycho-

analyst is taking a particular pattern of behavior as his unit—say the deprivation-desire-deed pattern in analyzing defences—which to the molar psychologist is quite a complex structure. I think that Skinner is in the position of a man who is asked for his suggestions at a research conference in the chemistry department, where they are tackling the problem of yeast assessment for bread-making. He says, "Find out all the physical properties of enzyme molecules and you'll have the answer." True, but only if old age is conquered in the meantime. Physics will some day perhaps be able to account for all chemical reactions, but today we have to evolve a selective approach via the chemical properties. It may not work, it may be necessary to get down to bedrock before a solid foundation can be laid; but it's well worth trying—and in chemistry it has paid off for a long time. Similarly in the social sciences, we can build up from the bricks of individual psychology or try to make a useful building with the large blocks of stones lying around with just a little chipping and perhaps a little more quarrying. And in psychology, we can work up from physiology or neurology, or tackle the issues more directly. Skinner is defending an approach to psychology which he rejects as an approach to psychopathology.

It is useless to complain that this is only a part of psychology; such an assumption is (a) extremely speculative and (b) even if true, no more effective than the argument that chemistry is only part of physics, which has hardly prevented chemists from discovering useful concepts at their own level. Skinner's extreme aversion to anything that looks like animistic talk leads him to miss the great wealth that can be mined from a logically sophisticated behavioristic analysis (not reduction—this is unnecessary and begs the philosophical question) of mental-state talk.

In particular, we can return to the two oversimplifications we made when beginning the discussion of "psychic states." A behavioristic analysis of "know" or "believe" has these two dimensions of flexibility (a better word than "vagueness," whose negative connotations spring from a failure to appreciate the necessity and utility of open texture in scientific language). How can this be useful in a science of behavior? How can it be anything but a drawback? The questions are immensely vain: for they presuppose that psychology is appropriate for or capable of answering such questions as What is knowledge? And, on the other hand, they are overly modest; for they presuppose that we have no idea how to shape the lists or estimate the probabilities.

Psychology should, of course, not be mainly concerned with epistemological questions, even though (a) its results may sometimes be relevant to them, and (b) the inspiration and description of research should not be independent of epistemological thought. Thus, psychologists concerned with "knowledge" characteristically ask such questions as these: What do seventh-grade math pupils know that sixth-graders do not? What does a man know about his childhood that he cannot immediately recall? Does a nondirective discussion result in the absorption of more knowledge than a lecture on the same topic? Does the teacher really know which of his pupils are the most intelligent? etc. There are endless experiments of this kind done, none of which fall into difficulties over the term "know"; though, of course, some others do, just as in physics some experiments are doomed from the beginning owing to faulty logical analysis. Only the crypto-philosophical psychologist inflates his experimental results into epistemological conclusions, and he is usually somewhat short on logical training. Even though some important questions are not properly answered by experiment—amongst them legal, literary, and logical ones—it is not necessary to reject as nonexperimental all problems involving mentalistic concepts, as Skinner does. For every man who can talk a language understands very well how to use these words, and can almost always tell a proper question from a nonsense-statement involving them. The fact that we can all use this language with great efficiency shows that one cannot judge the utility of a language by the test of whether it is reducible to a specific list of specific statements in the observation language. Nor even its scientific utility, for the language of many sciences has this feature. Thus for proper psychological questions, we shall, I suggest, no more frequently find ourselves in difficulty understanding or formulating a problem about knowledge or belief in ordinary language than if we invent some vaguely related and still imprecise (if still useful) language of our own. Inventing new terms is too frequently a substitute for analyzing the old ones and a move which only postpones the difficulties, for at some stage we try to relate the discoveries formulated in the new language to the problems formulated in the old. We can have our Taylor Scale on which we define a psychological concept of anxiety, but we can't avoid the question, "Does this give a good measure of anxiety as we ordinarily use the term?" Because if it doesn't, we can't use it to find out whether students are made more anxious by subjective than objective examinations,

or patients less anxious by piped music in the waiting-room. The background concept and theory cannot be ignored, though they can nearly always be improved; they have to be studied. A similar study of scientific terms reveals, below that surface sheen of operationism, the same open texture or flexibility. The difference lies in the purpose, not the nature, of the definitions employed, or employable except for some variations of degree (which are not all in one direction).

The language of psychoanalysis, in particular, is very open-textured; it is a first approach. Being so, it runs the risk of becoming empirically meaningless, a ritual form of mental alchemy. But the approach is fully justifiable; and it is as wrong to suggest that Freud should have pinned his terms down to infant neurology (CPCT, p. 302) or, by the "simple expedient of an operational definition," to physical and biological science (CPCT, p. 305), as it would be to insist that the founders of radio astronomy should have early said whether a radio star was a solid body or a region of space. They introduced the term as a name for the hypothesized origin of short-wave electromagnetic radiation. It now appears they were justified in using a name (i.e., spatial location is well supported), but we cannot yet tell exactly what radio stars (physically speaking) are, what the name stands for. Freud introduced the concept of the ego-ideal or superego as the hypothesized repository of the learned censoring activities of the personality. It is less certain that he was justified in using a name and quite unclear what, if any, physical reference it will have. It certainly need have no observable or measurable referent ("... the most unfortunate effect of all" those due to Freud's use of a "mental apparatus" approach [CPCT, p. 305]) to be respectable, though the hypothesis of its existence or operation must have observable consequences that are reliably identifiable. It was obvious to Freud that these consequences will have their neural counterpart if they exist and that the superego will thus, at least indirectly, have a neural counterpart; no one knew more than that. But everyone knew, or thought they knew, a great deal about conscience, anxiety, and guilt; and Freud discovered a great deal more, just as a psychologist might discover a great deal more about knowledge by an extension of the experiments described above. And just as the psychologist might quite accurately sum up some of his discoveries by distinguishing two types of knowledge (say kinesthetic and verbal), so Freud could distinguish

two stages in the development of the superego. Should we argue whether both types of knowledge exist or are real? We can, and we know how to do it in terms of correlation coefficients and χ^2 values—but it's an odd way to put the question. Does the superego exist? is it real? Well, we can deal with those questions, too, by pointing out certain undeniable features of behavior and arguing for their correlation and common subsumption under this heading. This doesn't show that the superego is observable, nor does it show that it is an explanatory fiction: Skinner's dichotomy is unsound. But—if it can be done—it justifies talking about the superego, which is as near as we can get to showing that it's real. It isn't real like a brain tumor, but it is real rather like an electric field; and it's certainly not unreal or a "fiction" or a "myth" like the aether—unless the arguments for it can be met, as those for the aether were met, on their own ground.

The other way of throwing over a theory, and a common one in the history of science, is to produce a better one. This Skinner would not be anxious to describe himself as doing. But his account of psychotherapy in SHB is an illuminating one, and it is theoretical. It is still far from being capable of dealing with the strange complexities of neurosis at an explanatory level, and it has no therapeutic success to support any claims for its practical efficiency. In fact, we have argued that while Skinner belabors Freud (somewhat unfairly, I think) for failing to give "an explicit treatment of behavior as a datum, of probability of response as the principal quantifiable property of behavior . . ." (CPCT, p. 304), in short, for being too theoretical, Skinner is himself a little too upset by the idea of explanations involving mental states to do them justice, even in his own operational terms. It is sometimes these errors of emphasis, rather than of fact, that lead us to abandon an approach. In his discussion of early psychological theories, Skinner comments on the simple reflex approach with these words: "It is neither plausible nor expedient to conceive of the organism as a complicated jack-in-the-box with a long list of tricks, each of which may be evoked by pressing the proper button" (SHB, p. 49). Reading Skinner, one sometimes wonders whether it is any more plausible or expedient to conceive of the organism as a complicated (but transparent) marionette with a long list of tricks, each of which may be evoked by pulling the proper string.