12-1976

On the Analytic Foundations of Behaviorism

Samuel Yaffe

*Western Michigan University*

---

**Follow this and additional works at:** [https://scholarworks.wmich.edu/masters_theses](https://scholarworks.wmich.edu/masters_theses)

Part of the *Psychology Commons*

---

**Recommended Citation**


[https://scholarworks.wmich.edu/masters_theses/2391](https://scholarworks.wmich.edu/masters_theses/2391)

---

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master’s Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the people who have contributed to making this thesis possible. I wish to thank Drs. Mountjoy, Michael and Falk for their guidance and assistance during the writing of this thesis, and especially to Dr. Mountjoy for his continual prompting to complete this project.

Samuel Yaffe
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeib Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
MASTERS THESIS

YAFFE, Samuel M.

ON THE ANALYTIC FOUNDATIONS OF BEHAVIORISM.

Western Michigan University, M.A., 1976
Psychology, general

Xerox University Microfilms, Ann Arbor, Michigan 48106
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>I BEHAVIORISM AND THE PHILOSOPHICAL \ DOCTRINE OF INTENTIONALITY</td>
<td>3</td>
</tr>
<tr>
<td>Intentionality</td>
<td>3</td>
</tr>
<tr>
<td>Intentionalism</td>
<td>11</td>
</tr>
<tr>
<td>A New Task</td>
<td>14</td>
</tr>
<tr>
<td>Summary</td>
<td>17</td>
</tr>
<tr>
<td>II A BEHAVIORAL APPROACH TO LANGUAGE</td>
<td>18</td>
</tr>
<tr>
<td>Interpretation and Verbal Behavior</td>
<td>18</td>
</tr>
<tr>
<td>Adequacy</td>
<td>22</td>
</tr>
<tr>
<td>Summary</td>
<td>24</td>
</tr>
<tr>
<td>III ...TO PURPOSE AND CONTENT</td>
<td>26</td>
</tr>
<tr>
<td>An Interpretation: Purpose</td>
<td>26</td>
</tr>
<tr>
<td>An Interpretation: Perception</td>
<td>39</td>
</tr>
<tr>
<td>Summary</td>
<td>44</td>
</tr>
<tr>
<td>IV AWARENESS AND BELIEF; A BEGINNING</td>
<td>45</td>
</tr>
<tr>
<td>Awareness</td>
<td>45</td>
</tr>
<tr>
<td>Beliefs</td>
<td>50</td>
</tr>
<tr>
<td>Summary</td>
<td>57</td>
</tr>
<tr>
<td>V ONTOLOGY</td>
<td>58</td>
</tr>
<tr>
<td>Hull's Argument</td>
<td>59</td>
</tr>
<tr>
<td>Hull's Bias Removed</td>
<td>63</td>
</tr>
<tr>
<td>Summary</td>
<td>65</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

(Continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>CONTENTS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI</td>
<td>OPERANT TERMINOLOGY</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Basics</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Some Niceties</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Ontology Again</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>89</td>
</tr>
<tr>
<td>VII</td>
<td>SUMMARY AND CONCLUSIONS</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>95</td>
</tr>
</tbody>
</table>
INTRODUCTION

The behavioral movement in psychology has been marked by energetic responses to repeated challenges of the general form "The behavioral approach cannot possibly account for phenomenon p, whatever its successes in other areas." Skinner's monumental *Verbal Behavior* (1957) is the result of a 23-year response to such a challenge. This paper is also a response to a set of related claims, made most notably by philosophers, to the effect that no theory such as Skinner's can possibly account for the set of phenomena frequently called intentional. I hope to show (i) that Skinner's theory can account for many, at least, of these intentional phenomena, and (ii) that his notion of operant behavior is to be recommended precisely because it is so relevant to the intentional (this is not to belittle the many other applications of operant theory to a wide variety of phenomena).

A major difficulty in achieving this goal derives from the fact that the challenge is couched in the technical language of modern philosophy, while the reply must be made in the technical language of operant psychology. This requires frequent translation of a sort, and the result runs the risk of being more confusing than enlightening. This is said not to shirk my own responsibility, but to spur the reader on as he tries to deal with the doubly confusing presentation in two complex and subtle languages. If experience is to be any guide, those readers who are familiar with one of the languages will tend to gloss over the difficulties and solutions presented.
in the other; such chauvinism will destroy any chance I might have of explaining myself. I can only hope that the required restraint and discipline will be rewarded by what I have to say.
CHAPTER I

Behaviorism and the Philosophical Doctrine of Intentionality

Intentionality

The concept of intentionality, or ones very similar to it, can be found in the medieval scholastic philosophers (e.g., Duns Scotus; c.f., Copleston, 1950), and some have argued that the beginnings of this view are to be found in Aristotle (e.g., Boring, 1950). These precursors, however, may be regarded as being of little but scholarly interest, since the conception takes on its distinctly modern tone in the work of the psychologist-philosopher Franz Brentano (1838-1917), a contemporary of Wilhelm M. Wundt (1832-1920).

Brentano's influence on psychology can be estimated in one way by naming some of the psychologists he influenced more or less directly: Carl Stumpf, Christian von Ehrenfels, Edmund Husserl, Alexius Meineng, Oswald Kulpe, and Sigmund Freud. Another estimate can be made by noting that his students began what Boring calls "experimental phenomenology" (boring, 1950), as well as the school of Gestalt psychology. An estimate which is of particular relevance here begins with Boring's view that "...German psychology at the end of the last century..." revolved about "two foci:" set systems, and content systems, "the one represented by Brentano, the other by Wundt" (Boring, 1950). Boring suggests that

...act psychology could have developed into the modern dynamic psychology of behavior and purposive conduct without there ever having been a psychology of content (Boring, 1950, p. 439).
A union of the two foci was attempted by Kulpe and August Messer; Boring characterizes this attempted "marriage" of act and content as one which was "...never destined to persist..." (Boring, 1950). This historical digression is relevant because one way to characterize the intentionalist thesis is that behaviorism can never get "...from motion and matter to content and purpose...", and again that it can never account for both content and purpose, which are seen by the intentionalist as in a state of "interpenetration" (Dennett, 1969). The intentionalist, then, sees Boring's marriage as a fact which persists in the world, whether or not psychologists have been able to sustain the union in their theories. The reader is here forewarned that Skinner's theory will be analyzed as one which involves in its most basic conceptualization a union of purpose and content, which the psychologist might more readily recognize as goal orientedness and stimulus control.

The modern tradition of intentionality begins with Brentano's *Psychologie vom empirischen Standpunkt*, and his contribution is summarized in two short passages. He takes it as given that

> The data of our consciousness make up a world which, taken in its entirety, falls into two great classes, the class of physical and the class of mental phenomena (Brentano, 1874, emphasis in original).

For various reasons, Brentano took great pains to discriminate clearly and precisely between these two classes, and he arrived at the following critical differences:

> ...psychologists of an earlier period have already directed attention to a particular affinity and analogy which exists among all mental phenomena, while the physical do not share in it. Every mental phenomenon is characterized by what the scholastics of the Middle Ages called the intentional (and
also mental) inexistence (Inexistenz) of an object (Gegenstand), and what we could call, although in not entirely unambiguous terms, the reference to a content, a direction upon an object (by which we are not to understand a reality in this case), or an immanent objectivity. Each one includes something as object within itself, although not always in the same way (Brentano, 1874, p. 46).

Brentano summarized his thesis by saying that this "...intentional inexistence is exclusively characteristic of mental phenomena. No physical phenomenon manifests anything similar..." (Brentano, 1874). The implicit challenge to behavioral theories is clear: behaviorism strives to measure and define its variables in physical terms, which will inevitably preclude its accounting for the mental life of men, because mental phenomena have an essentially different characteristic, that is not had by any physical phenomena. Before presenting a behavioral reply to this challenge, Brentano's language, rather confusing to modern readers, must be explained.

In what ways, then, does a mental phenomenon include "something within itself"? Brentano provides two answers: by "reference to a content" and by "a direction upon an object." By "a direction upon an object," Brentano seems to indicate such phenomena as wanting, hoping, and seeing. One cannot want without wanting something, hope without hoping for something, or see without seeing something; in each case, the "something" is the object the mental phenomenon is directed upon. By calling these objects "intentionally inexisten," Brentano is simply pointing out that, while they are directed upon by the phenomenon, they needn't exist in the normal sense of the term. That is, the fact that one wants something, hopes for something, or sees something, does...
not mean that the something truly exists to be hoped for, wanted, or seen. Skinner expresses a very similar notion when he says that "...seeing does not imply something seen" (1963a, p. 234). This is most certainly in contrast to such physical phenomena as collisions, where, if A collides with B, both must certainly exist. Since the intentional object does not truly exist, Brentano infers that is contained within the mental act itself.

By "reference to a content" Brentano indicates such things as beliefs, and some kinds of hopes. Often enough, one believes that some proposition is true, or hopes that it is true, although this in no way guarantees that the proposition is in fact true. Moreover, one can hope to see the New World without hoping to see America, and this shows that the object upon which intentional acts are directed is one with, as it were, a particular content. In contrast, one cannot blow up the New World without blowing up America.

Indeed, this non-substitutability of terms which denote identical objects is a key feature of contemporary discussions of intentionality. Philosophers generally shift to the linguistic level, from talk about phenomena, as Brentano's, to talk about talk about phenomena. Single quotes are used to refer to or name the expression within the quotes, so we can say that 'x wants y' does not logically entail 'y exists.' And 'x wants y' and 'y is identical with z' do not conjointly entail that 'x wants z.' Finally, 'I believe that p,' where 'p' names a proposition, does not entail that p is true, nor does 'I believe that there are unicorns' entail 'There are unicorns,' nor 'I believe that there are one-horned animals.' Thus, there are certain logical oddities about
sentences describing intentional phenomena, and these oddities—the non-substitutability of identicals and the invalidity of existential inferences—parallel Brentano’s ways in which the intentional act and its object are related.

Of course, neither Brentano’s description nor any of the foregoing examples suffice to tell anyone exactly what intentionality (the property of being intentional) is. Indeed, the precise specification of the class of intentional sentences remains a philosophical problem. However, if the reader has gained some weak tendency to discriminatively respond to intentional sentences from these examples, he will be better able to follow the following discussion of two philosophers’ attempts to offer necessary and sufficient criteria which fit their intuitions concerning Brentano’s insight.

Chisholm has formulated four independent criteria, which he takes as a working definition of intentionality, although he is aware that they are not quite enough (Chisolm, 1957, pp. 170-171). They are, in essence, as follows:

(C1) A simple declarative sentence S is intentional if
(i) S uses a name or description N, and
(ii) neither S nor not-S imply either that there truly is an N, or that there is not.

(C2) If S
(i) contains a clause which expresses a proposition P, and
(ii) neither S nor not-S imply P, nor not-P, then S is intentional.

(C3) If
(i) I is a sentence expressing the proposition that one name or description N1 names the same thing as another, N2, and
(ii) one sentence S1 differs from another S2 only in
having N1 where S2 has N2, and
(iii) S1 and I together do not imply S2, then
S1 and S2 are intentional.

(C4) If S is formed by connecting more than one
simple sentence by such expressions as 'if-then,'
'and,' 'or,' and 'because,' then S is intentional
if and only if at least one of its simple sentence
components is intentional by C1-C3.

For example, 'Joe hates beggars' is intentional by C1, since even
though there are beggars, Joe's hating them or not hating them does
not imply that there are or are not beggars. 'Joe is in bed,' on
the other hand, does imply that there is a bed, so it isn't inten­
tional by C1. Using C2, 'Skinner holds that there are no minds' is
intentional, since the truth or falsity of 'there are no minds' does not
follow from Skinner's holding it to be true or not. An example of
C3 is 'Joe is contemplating the President's evils,' for 'the President'
and 'Nixon' may name the same thing, but Joe may not know this, so it
does not follow that he is contemplating Nixon's evils.

That C1-C4 do not exactly capture Brentano's conception is shown
by the fact that 'Joe looked for the morning star' does not entail
'Joe looked for Venus,' although 'the morning star' and 'Venus' name
the same thing. Thus, although it is not obviously about mental phe­
nomena in Brentano's sense, as he took such phenomena to be "private
events" (Brentano, 1874, p. 37), 'Joe looked for the morning star' is
intentional by C3. As Dennett says, "the usual line" philosophers
take at this point: "...is to reconstrue 'mental' rather broadly, or
to replace it with 'psychological'..." (Dennett, 1969, p. 23), rather
than abandon C3. Dennett also says that the usual line is to "...accept
such sentences as falling within Brentano's notion..." of intentionality (1969, p. 23). This reinterpretation of intentionality as the mark of the "psychological" may be unexceptionable, but it is clear that this move constitutes a revision of Brentano's view, still, it may retain his insight into a cleavage between two classes of phenomena, the one having to do with physical events, and the other with distinctly different ones.

Dennett's reason for holding that this broader interpretation of intentionality falls within Brentano's notion is worth looking into. He considers a group of what he calls "mongrel terms," including 'looked for.' Consider:

(D1) Tom looked for the cat.
(D2) Tom perceived the cat.
(D3) Tom recognized the cat.
(D4) Tom discovered the cat.
(D5) Tom hallucinated the cat.

According to Dennett, D1-D5 are intuitively psychological, and thus intentional, although D1-D4 imply (for him) the existence of the cat, while D5 implies its non-existence. These, and a host of other examples, he takes as "...clearly mongrel terms, part mental or psychological, part contextual or epistemic..." (1969, p. 24). His approach is to take sentences with such terms as intentional in view of their intentional "implications" (1969, p. 24f). For example, D2 is taken to imply:

(D6) It seemed to Tom that he perceived the cat.

which is intentional by C2. Dennett would undoubtedly claim that D3-D5 also imply D6, and are thereby also intentional; D1 might be taken
to imply 'Tom hoped that he would find the cat,' thus insuring its intentionality. Chisolm makes the same move in his discussion of perception:

When we use perception words propositionally, our sentences display the third of the above marks of intentionality. I may see that John is the man in the corner and John may be someone who is ill; but I do not now see that John is someone who is ill. Perception sentences... entail sentences about taking and assuming. And sentences about taking and assuming display the second of the above marks of intentionality (Chisolm, 1957, pp. 170-171).

Taking such sentences as D1-D5 as intentional is unobjectionable; they are indeed intuitively psychological, and Dennett's mongrel term analysis reveals psychological aspects in the ordinary sense of the term. However, the mongrel term approach is merely intuitive, however accurate, and the implication criterion suggested by Dennett and Chisolm will not do. Sentences like 'Joe hit a home run' are intuitively non-intentional, but home runs are not physical events, although they involve physical events. Something is a home run only according to certain rules of the game of baseball: the umpire must judge it a fair ball, that Joe touched each base as he ran around the diamond, etc. In short, 'Joe hit a home run' implies that someone (the umpires) took it that certain physical events took place, which, according to Chisolm, makes 'Joe hit a home run' intentional.

There may be some reason to take such sentences as intentional after all, but if the implication criterion is the only reason, the situation becomes untenable. For any sentence, apparently as unintentional as you like, becomes intentional. Thus, 'This is a table' implies that the speaker takes it to be a table. On the other hand,
if the implication or mongrel term criterion is abandoned, D5 at least presents a decided problem. For D5, contrary to Dennett, neither implies that there is nor that there is not a cat. I may hallucinate a cat, even if it is in the next room, or even if it is before me and, say, my eyes are closed.

The upshot of this discussion is that C1-C4 provide very good, but imperfect criteria for discriminating intentional from non-intentional sentences. Chisolm's criteria are helpful, of course, in that they do provide very good guidelines, and help sharpen one's intuitions so that analyses such as Dennett's of mongrel terms can be seen to make sense. So even without perfect definitions, the distinction between intentional and non-intentional sentences may be taken to mark a real difference in the world.

**Intentionalism**

Having been led through this philosophical maze to a rough distinction between intentional and other sentences, and thereby a corresponding, implicit distinction between the intentional and other phenomena described by such sentences, the reader may well wonder how the distinction is of any moment for psychology. Hull sheds some light on the problem by referring to "...non-causal, non-behavioral, non-linguistic features..." of phenomena like perception, and calling these features intentional (Hull, 1973, p. 65). Obviously, if there are non-causal, non-behavioral, non-linguistic aspects of human activity, then Skinner's program of radical behaviorism is bound to fall short
of the mark, for his account, which eschews non-causal, non-behavioral, non-linguistic terms, will leave such features unexplained. And if this something left out includes a significant part of what we want to explain about human activity, then the discerning psychologist will want to look elsewhere for a complete theory of human behavior.

The intentionalist challenge can be brought into sharper focus by returning to intentional sentences as defined by C1-C4. While there may be dubious cases, it is clear that the behaviorist cannot reject all such sentences as meaningless. On the other hand, he cannot accept them as formulated either, for many of them use terms which are entirely unacceptable in that they do not clearly specify any physical dimensions. For example, in

(a) He hates dogs.
(b) He is daydreaming.
(c) He believes in Catholicism.

there is no obvious physical dimension(s) specified by the emphasized words. Still, the sentences describe very real situations. The challenge is to describe the same facts exactly, without adding or deleting anything at all, and entirely in the physicalistic language of reflexes of type S, type R, and the variables of which those reflexes are a function.

The difficulty of the challenge, as well as the ubiquity of the intentional, can be seen by trying to meet the challenge in a particular case. Let us take (a) above. One large class of overt acts is ruled out because it involves substituting one intentional expression for another. Thus, 'He beats dogs' is as intentional as (a), for 'beats'
and 'strikes hard' may designate the same physical movements, but not
the same action, since striking hard may come from a reflexive move
of defense, which is not beating; and, of course, one may hate dogs
and still never beat them. Indeed, one may truly hate dogs without
ever emitting any overt behavior as a result, which means that any
behavioral translation must appeal to some property other than overt
behavior as a result, which has been the anti-behaviorist charge for
many years. The behaviorist is thus chased from his stronghold at
the outset.

Another approach, which is often tried, is the dispositional one.
The hating of dogs is translated into the agent's disposition to be­
have in certain ways. The problem here is the specification of the
behaviors he who hates is disposed to emit. We can choose from such
things as saying he hates dogs, hits them, avoids them, and so on, but
in each case it is easy to imagine him hating dogs and not doing any
of these things, or doing them but not hating them.

There is a large body of literature containing other examples of
putative translations which either resort to more subtle intentional
idioms, or are simply inadequate translations altogether (e.g., Dennett,
1969; Chisolv, 1957; Sellars and Chisolv, 1958). The situation has led
Davidson to say

Why are we willing (as I assume we are) to abandon the
attempt to give explicit definitions of mental concepts
in terms of behavioral ones? Not, surely, just because
all actual tries are conspicuously inadequate. Rather, it
is because we are persuaded, as we are in the case of so
many other forms of definitional reductionism (naturalism
in ethics, instrumentalism and operationalism in the
sciences, the causal theory of meaning, phenomenalism,
and so on—the catalogue of philosophy's defeats), that there is system in the failures (Davidson, 1970, p. 91).

I think it is important to see in this passage sincere regret on Davidson's part, that one is not deluded into thinking that the many philosophers who share his view do so out of some vendetta against behaviorism. Their conclusions are often drawn reluctantly, after a great deal of resistive analysis. They simply have not been able to find an acceptable way out.

A New Task

A clue to the origin of Davidson's "system in the failures" is found earlier in his paper.

Do we declaring that there are no (strict) psychophysical laws, poach on the empirical preserves of science—a form of hubris against which philosophers are often warned? (Davidson, 1970, p. 90).

Davidson concludes that his inquiry treads a vague line between philosophy and science, and he declares that "Where there are no fixed boundaries, only the timid never risk trespass" (Davidson, 1970). I have no quarrel with his decision; I only suggest that travelers should learn the language(s) spoken where they go, especially if they have reason to suspect that they are trespassing. It may be that the system in the conspicuous failures which lie strewn about the philosophical literature is the philosophers' systematic refusal to utilize experimentally derived concepts. It is my hope in this paper to provide a few conspicuous successes by relying on the system of concepts developed by Skinner and his fellow operant psychologists.
A limitation of this approach—and, perhaps, an explanation of philosophers' avoidance of it—is that it ties one's wagon rather closely to the fortunes and misfortunes that may befall Skinner's conceptualization. Thus, to the extent that successes are obtained they are contingent upon continuing confirmation of the empirical concepts utilized. This is both cause for optimism and for caution. Insofar as certain concepts provide a key to a solution, those concepts may gain in respectability and adherence; insofar as they meet with empirical disconfirmation, they render the present analysis less conclusive. However, even in the latter eventuality, a success will count this much: it will serve to still the clamor that behaviorism is impossible. A second limitation stems from the number of possible intentional idioms (which is likely countably infinite): we can analyze only a few, and are bound to omit many that may seem crucial to some. There is nothing for this second problem; we can only empathize with Lincoln, who also despaired of pleasing all of the people all of the time.

Still, the approach to be followed promises a benefit rarely anticipated by most philosophical inquiries. As typically formulated, intentionalism presents a dilemma: either Brentano was right, or there is no interesting distinction between psychological and other phenomena. Taking the former horn, some philosophers draw conclusions other than the simple impossibility of a reductionist, physical account. Dennett (1969) and Davidson (1970), in different ways, argue that a physical science of behavior may be possible, but that such a science
must coexist with the traditional intentional "theory of mind" in a way that precludes any account of the latter in terms of the former.

It is difficult to see how this sort of side-by-side approach would work, for verbal behavior, including the production of intentional idioms, would certainly seem to be an important, if not essential, realm of physical behaviors to be explained by any physical science of behavior; and, it is clear that an account of the production of some such idiom in terms of Skinner's approach will provide, ipso facto, an account of the intentional phenomenon described by _i_. Hull (1973), perhaps seeing the inadequacy of this side-by-side approach, argues, in effect, that a philosophy of psychology such as Skinner's is actually inconsistent with its presumed goal of explaining the ordinary (intentional) phenomena that occasion the enterprise of psychology at the outset! We shall see that, by utilizing Skinner's notion of the operant class, there need not be any dilemma at all. For the very phenomena which occasion the use of intentional idioms will be found to be phenomena of operant behavior. And it is rather clear (though there are dissenters here) that the functions describing the laws of operant behavior are easily distinguished from those describing either respondent behavior or the purely physical behavior of mass and energy. Within the limits set down above, then, it follows that Brentano's instinct was correct: there is a distinction to be made between psychological (or operant) phenomena and other phenomena, which is discriminated linguistically in ordinary language. However, the distinction is between relations among familiar entities, not between basic kinds
of entities requiring radically different kinds of language. Once we get clear about the way operant analysis affords an analysis of the intentional into its physical parts, the dilemma disappears, and the remarkable conclusions of Davidson, Dennett and Hull can be seen as unnecessary.

Summary

We have traced the history of Brentano's conception of a fundamental distinction between the mental and the physical. Some progress was made in clarifying his notion of intentionality, and a more modern scheme of defining this property was discussed. It has been shown how philosophers see in Brentano's idea a fundamental problem in the methodology of psychology, and a few examples of how philosophers have failed to reply adequately to this challenge have been discussed. Three different conclusions have been drawn from Brentano's problem, none of which seem to fit into the program of behavioral psychology. However, Davidson's discussion of these results has given us an insight into a possible solution that has not been fully examined. Accordingly, we must begin again, with a new approach to language itself, an approach based upon the empirical results of operant psychology.
CHAPTER II

A Behavioral Approach to Language

Interpretation and Verbal Behavior

Since our characterization of intentionality is linguistic, it is appropriate that our discussion should begin with Skinner's account of verbal behavior, which he characterizes as an "interpretation" (Skinner, 1957, p. 4). Interpretation, as Skinner uses the term, is an activity or the result thereof in which "...you make plausible guesses at the variables which are probably operating in a given case" (Evans, 1968, p. 105). An example is the physicist musing over the cooling of his morning tea, guessing that breezes from the window, the conductivity of his cup, reducing the volume by drinking, etc., are the variables which are probably operating, according to laws discovered in the laboratory, to reduce the temperature. Of course, such interpretations do not prove that the phenomena occur as interpreted: some important variable might not be considered, some parameter values might be set too low or too high, and research may reveal new laws. Nevertheless, unless he appeals to some implausible factor, say conductivity in his spoon that the known laws of physics preclude, or a kind of metal as yet unknown, his interpretation renders plausible the claim that the phenomenon is identical with a physical one.

Similarly, an operant interpretation of a given case of verbal behavior, if plausible, renders plausible the claim that the verbal
behavior is identical with an operant phenomenon. Of course, we are limited to interpretation. Without a technology to test such guesses, a technology which is almost certainly quite far off, we must rest with some sort of intuitive test of plausibility. Actually, the procedure is quite similar to a kind of analysis typical in philosophy, so philosophers should find themselves quite at home. However, this procedure necessarily limits us to cases in which our intuitions are clear.

It will become evident that operant interpretations of intentional language tread closely to the traditional conception of meaning. As Day (1969) points out, Skinner is at one with Wittgenstein in virtually rejecting the concept of meaning; more precisely, both look for those characteristics of a situation which lead a verbal community to approve of, or reinforce, a given utterance in the situation. Wittgenstein (1953) has pointed to one difficulty with this approach in his distinction between symptoms and criteria. For example, "Fire!" may be reinforced in the presence of smoke, but there is surely something interestingly different when the presence of flames is a key factor. There is another difficulty: especially in the case of first person statements about private events, we shall need to distinguish the speaker's "meaning" from the symptoms or criteria leading to communal reinforcement. Finally, a difference must be noted between a speaker's "meaning" and that "taken" by the listener. Wittgenstein found these problems particularly thorny, and the reason, I suspect, is that as a philosopher, he refused to lean on any but the most evident contingent
truths. By relying on the basic operant phenomena of discriminative behavior, generalized conditioned reinforcers, etc., we may hope to deal more effectively with these problems.

First, the speaker versus the listener. The speaker's meaning is to be analyzed in terms of the variables of which his utterance is a function; that is, once the functions are understood in detail, one is no longer prone to ask "Yes, but what does he mean?" Another approach to the speaker, of course, lies in the listener's responses that the speaker will find acceptable and those he will find unacceptable. Now, the utterance will serve as a stimulus for the listener, whose behavior is also a function of other variables. We may look at the listener's response as a function of all of those variables, or only the difference made by the utterance itself; either way, we have an answer to one sense of the question "How did the listener take that utterance?"

The symptom-criterion distinction is handled on two levels. In one way, there is no distinction. For the reinforcement of "Fire!" as a function of the presence of smoke simply shows us another use of that utterance: it is proper to call "Fire!" when you see smoke billowing out of a building, regardless of whether or not there are any flames involved. There is no need to assume that the speaker was guided by the rule 'Where there's smoke there's fire,' although he may have been. In some contexts we simply reinforce "Fire!" while we have never (or rarely) reinforced "Smoke!"

Still, criteria must be distinguished from mere symptoms, if for no other reason than that the verbal community does so. Wittgenstein's
observation that questions about such matters must come to an end, and rather abruptly in many cases, shows that our practices implicitly sanction the distinction; sooner or later we just say that this is called a bleep: that's the way we speak. And that same observation shows where the line is drawn. For example, one might imagine that a fair ball is one that is hit within certain boundaries, and we certainly infer fairness from evidence regarding boundaries. However, the fact is that a fair ball is one that is so designated by the umpire. Instant replays may show that the ball was hit within the boundaries, and sportscasters may even say correctly enough that "It really was fair," and the umpire might actually have been mistaken about where the ball went. However, the umpire cannot possibly be mistaken about its having been foul, for the practices of professional baseball are such that he is the final arbiter on such questions. To be sure, practices do change, and we might come to distinguish between the "actuality" and the "call". However, that we have not done so can be seen by asking how many hits an inveterate foul ball hitter got in a given year; we do not count the hits called foul that went into the stands, even if tapes show conclusively that some of them were within the proper boundaries. And the point is only that there is an entity, whether foul balls or balls called foul, the actuality of which is a function of the umpires' behavior.

The baseball example is unusual in two ways. The criterion is the verbal behavior of a given person(s). Many of our "language games," as Wittgenstein called them, involve non-verbal criteria, and their effect
upon a community rather than a specific individual(s). This latter fact is a crucial one, for it leads to a statistical and rather arbitrary specification of "the meaning" of an expression. Arbitrary in that we can consider a verbal community as containing all speakers of the language, only those for whom it is their first language, educated speakers only, etc. And statistical in that, for any given verbal community, some degree of variation is to be expected in the variables which lead given members to cease doubting or arguing.

In brief summary, we have distinguished three levels on which given verbal behaviors might be interpreted: the variables responsible for the speaker's utterance, the variables responsible for the listener's response, and the variables responsible for the verbal community's reinforcing the utterance. These correspond roughly to traditional notions of the speaker's meaning, the listener's understanding of the speaker, and the standard meaning of the utterance, respectively. The correspondence is only rough because, for example, the speaker's behavior might be understood only by taking into account, say, the listener as stimulus, which may have nothing to do with his meaning. For now, the point is that the traditional views differ in their referents, and that the operant view allows of any of several analyses of "meaning," each corresponding, however roughly, to one traditional view.

Adequacy

Before diving into specific interpretations, some standards are needed in rough form to distinguish adequate from inadequate interpretations. Intentional expressions must be avoided, of course, and
functional relations unknown to the literature of science cannot be allowed. Adequate interpretations may involve such statements as 'Intentional description I designates x, Behavioral description B designates y, and x is identical with y.' Many questions can be raised concerning any statement of identity, but it will be enough to show plausibly the sort of identity that is supposed to hold between the cooling of a cup of tea and the decline of its mean kinetic energy.

A particularly trucludent problem derives from the continuum of specificity that is possible in interpretation, and leads quickly to the kinds of properties or events that will be allowed in the description of variables. Skinner moved to a class definition of stimuli and responses in 1935 (Skinner, 1935a), and he recognized then that a consequence of this approach is that "...the existence of a popular term does create some presumption in favor of the existence of a corresponding experimentally real concept..." (Skinner, 1935a, p. 363). However, he also admonished that...

...this does not free us from the necessity of defining the class and of demonstrating the reality if the term is to be used for scientific purposes. It has still to be shown that most of the terms borrowed from the popular vocabulary are validly descriptive—that they lead to consistent and reproducible experimentation. We cannot legitimately assume that "riding a bicycle," "seeing one's friend," or "heartbreak" are responses in any scientific sense (Skinner, 1935a, p. 363).

The point is unexceptionable, and provides a gentle reminder that radical behaviorism involves an analysis of the parts of intentional phenomena. That is, it is unexceptionable in the experimental context, but what of the non-experimental but empirical activity of interpretation? Much of Skinner's interpretive works do not follow this admonition;
and, of course, they could hardly do so, for they would then be experimental and not interpretive at all. There is only one possible solution: we must rely on our intuitions of the plausibility of the validity of our terms, and make them only as specific as is necessary to satisfy our intuitions. For example, there will be no need to analyze the distinguishing features of dogs; it will be acceptable to speak of a dog as a stimulus, and dogs as a stimulus class, for most speakers.

A closely related issue concerns the breadth demanded of the analysis. On the one hand, we can hardly reject every distinction available in ordinary language; and, on the other hand, we cannot expect that every distinction enshrined in our vocabulary reflects a real difference among non-verbal phenomena. Reliability of usage is critical; that is, when virtually all speakers agree upon which things are bleeps and which are not, it will be difficult to deny the distinction between bleeps and non-bleeps in the world. Finally, this problem and many others can be minimized by staying close to relatively clear cases.

Summary

We have explored the behavioral approach to language, and characterized interpretation, the activity of making plausible guesses concerning the variables responsible for a given utterance in a given situation. It was seen that the interpretation of a given expression is akin to an analysis of its meaning, and several kinds of meaning were distinguished. The central role of intuition in judging the adequacy of interpretations was stressed, and the need for clear cases was
seen as a corollary. The next step is to illustrate the method by utilizing it in the interpretation of intentional idioms.
...To Purpose and Content

In a particularly insightful work on the problem of intentionality, Dennett described the situation in this way:

...the evidence comes in that we can neither do without the Intentional, nor cleave to it alone, for there are signs that the possibility is remote of a successful non-Intentional behaviourist psychology... (1969, p. 39).

Dennett has, obviously, considered the behavioral approach and found it wanting, and has also considered the possibility of a science of psychology couched entirely in intentional terms. He rejects the intentional approach, because he finds it leading to a drastic revision of the entire logic of science, and, as he sees no other alternatives in the literature, he characterizes the situation as a dilemma. He is not without hope, however, and the rest of his book is devoted to developing a third approach. And he sets his own criterion thusly:

The task of avoiding the dilemma of Intentionality is the task of somehow getting from motion and matter to content and purpose—and back (Dennett, 1969, p. 40).

His analysis is convincing enough to this writer to foster a course in which the defense of the behaviorist position begins with interpretations of purpose and content.

An Interpretation: Purpose

English is an incredibly rich language, with many idioms available to denote the purposive or, in the ordinary sense, intentional aspects
of behavior. Although certain subtle contextual variables may favor one over another in a particular case, the following schema are roughly equivalent (where \( x \) names an individual, \( y \) an act or response, and \( z \) an event):

(a) \( x \) did \( y \) in order to get \( z \)
(b) \( x \)'s goal in doing \( y \) was \( z \)
(c) \( x \)'s intention in doing \( y \) was \( z \)
(d) In doing \( y \), \( x \) meant to affect \( z \)
(e) \( x \)'s purpose in doing \( y \) was \( z \)
(f) \( z \)'s reason for doing \( y \) was \( z \)

There are no doubt many others, and these may not be equivalent in all contexts, although there certainly are contexts in which all would be approximately equally appropriate. And they are all intentional by Cl, since the existence of \( z \) in each case is not entailed by them, nor by their contradictories. Taking a specific instantiation of (a),

(1) The rat pressed the bar in order to get a food pellet.

we find what may be called the paradigm of operant behavior. And small wonder, since Skinner (1963b) suggests that his concept of reinforcement takes the place traditionally taken by the concept of purpose.

Of course, (1) is not likely to be spoken outside the laboratory, and it is not likely to be reinforced even there. Rather, laboratory practices require some such thing as

(2) The rat moved in such a way as to result in the depression of the bar, in part at least because past movements which depressed the bar were reinforced by a food pellet.

where 'because' is a "causal" or functional term. This suggests a quite general interpretation of
(3) x did y in order to affect z

as

(4) A causal (or functional) factor in x’s doing y was a history in which a response(s) belonging to the same class as y were reinforced by an event(s) belonging to the same class as z.

This suggestion boils down to the claim that the propriety of sentences fitting schema (3) is in fact accepted by the English verbal community on the basis of evidence for the truth of (4), and that conclusive evidence for the truth of (4) would be conclusive evidence for the propriety of sentences fitting schema (3). This may seem far-fetched, since most members of the community would not even understand (4); but, then, many addicts can identify opiates by carrying out essentially biological observations, in spite of their ignorance of the relevant science. The point is not that sentences fitting schema (3) are adjudicated by first ascertaining the truth of (4); rather, the suggestion is that the observations made in adjudicating the one are essentially the same as those made in adjudicating the other. There will be differences, of course; scientific practices are both more exacting and thoroughgoing than ordinary ones. An analogy may be helpful, and it is important that this be made clear, for we are dealing again with the nature of interpretation. One mechanic judges the car to be out of tune by listening to it idle, another by using a timing light, and a third by using an oscilloscope; their observations differ in many ways, but the same condition of the engine is responsible for all three judgements. In addition, each may describe his findings in terms of his observational method, but each description is right or
wrong as a function of the actual condition of the engine. The claimed relationship between (3) and (4) is that the former is true if and only if the latter is true, even though the speaker of (3) cannot understand (4), and even though he cannot understand that he is describing the same phenomenon as the psychologist who utters (4). Still, our best description of the claimed relationship is to say that, if it were necessary, any fact describable by schema (3) could be described by (4), with the only changes in meaning being those that inevitably arise when one moves to a scientific description. Or, again, we could say that (3) and (4) are perfectly co-extensive, applying or not in exactly the same cases, and for exactly the same reasons.

Let us try a specific case:

(5) Mo deposited a coin and pushed the button in order to get a soda.

For purposes of analysis, (5) can be regarded as a conjunction of

(6) Mo deposited a coin and pushed the button

and

(7) He did so in order to get a soda.

Many philosophers will claim that (7) entails that Mo had various beliefs or knew certain things, while interpretation of (7) in terms of (4) does not. Now, it cannot be denied that beliefs and/or knowledge are often involved. Nevertheless, that nothing more than a relevant history of reinforcement is necessarily involved can be seen by removing the rich background assumed in the typical case. Assume that Mo is young, retarded, an aborigine, or anyone lacking in familiarity with vending machines. Observing Mo's behavior, Op may utter (7), and Int
may object with "My dear Op, that's perfectly absurd; he doesn't understand the first thing about vending machines." This is a proper objection insofar as it claims that Mo couldn't have had any reinforcing experiences with such behavior. Op may reply "Oh yes he does; he was taught only yesterday, after a sweaty workout... just like the one he's just finished." Op has two crucial bits of relevant information Int lacks: he knows that Mo has been reinforced with a cold soda for operating the machine, and that a relevant deprivation state existed to make likely responses maintained by cold drinks. That the experience of having been reinforced for such behavior is crucial to both Int's objection and Op's reply shows the crucial role played by such experiences. Now, some may hold that, having had such an experience, Mo should be described as having some relevant knowledge or belief, even if Mo happens to be a trained chimpanzee. Insofar as beliefs or knowledge are entailed by the reinforcing experience or trial, it must be that we ascribe said states on the basis of evidence for a relevant history of reinforcement. Thus, in this basic case at least, beliefs and knowledge are either not involved at all, or are superfluous additions to the basic purposive nature of the behavior.

In one case where Mo is not lacking in vending machine experience, the relevance of reinforcement history may be indicated only by a secondary variable, like generalized conditioned reinforcers. Int might object again: "But this is the tenth soda he's gotten in half as many minutes!" A specific reply for Op might be "Well, then, he must have lost the draw, so he's getting sodas for all the others,"
which would mean that obtaining the soda was one reinforcer in a behavior chain. Indeed, the concept of chaining has to be invoked already for (5), since 'deposited a coin' and 'pushed the button' are elliptical for 'moving thusly to get a coin in the slot' and 'Moving thusly to get the button depressed.' According to the operant analysis, the disappearance of the coin in the slot and the depression are discriminative stimuli (and, perhaps, conditioned reinforcers), which are essential components of the chain that leads to the soda.

One objection to (4) as an analysis of purpose might be made in terms of novel responses. Suppose that Joe's purpose in throwing a rock is to break Int's window, and that he has never before broken that window. We cannot interpret

(8) Joe threw the rock in order to break Int's window.

in terms of (4), for we would get

(9) A factor in Joe's throwing the rock was a history in which rock throwings by Joe were reinforced by breakings of Int's window.

and (9) is false ex hypothesis, for Joe has never broken Int's window before. Indeed, the case can be made stronger by further assuming that Joe has never broken any windows, has never thrown a rock, has never thrown anything, etc.

However, it is clear that Joe must have had some prior experience relevant to the case. Had he never thrown anything, never broken anything, nor observed any such behaviors or consequences, nor read about them, one would be tempted to conclude that he was, in fact, a puppet under the control of someone who had had the relevant experience, or some other equally fantastic conclusion. The point is that, behaviorism
aside, we demand of any account of human behavior that particular acts
be related in some way to the agent's history or experience. Thus,
it is not enough to object to the analysis of (8) in terms of (4) that
Joe lacked certain experiences; one must also specify the relevant his­
tory Joe did not lack. It is among the antecedents of Joe's behavior
that we must look for an interpretation of his purpose in terms of (4).

Suppose, for example, that Joe had broken other windows by throw­
ing rocks, and that he was dared to break Int's by his friends. Under
this sort of multiple control, Joe's behavior could correctly be said
to have, as its purpose, avoiding the aversive consequences "implied?"
as it were, in the dare, obtaining the positive consequences similarly
implied, acting in accord with the dare, or, what is the same thing,
doing what he was dared to do and, therefore, to break Int's window.
The problem, then, is to show how an interpretaion in terms of (4)
will yield these as purposes in this particular case.

The first pair are easily accounted for on the bare assumption
that he has been dared in the past. Thus, if in the past, behaving
in accord with dares has led to certain positive reinforcers (peer
approval, etc.) and the present behavior is in part a function of this
history, obtaining positive reinforcement of this sort is a purpose of
this behavior by a straightforward application of (4). And, if he
has had a history of successful avoidance in the context of similar
verbal behavior (i.e., threatening tones), and the present behavior
is in part a function of this history, (4) yields avoiding punishment
of this sort as a purpose. (There might be objections to this use of
'acting in accord with a dare.' But the problems thus raised have nothing to do with the question of the behavior's purpose; they regard the more general question of verbal behavior and its efficiency as stimulus. Intentionality figures here only in the most general way, and does not affect the narrower question of purpose.)

Acting in accord with the dare appears to present an unusual case only if we insist on seeing here a description of a response and nothing else. In fact, however, it is quite plausible to suppose that having so acted will have been a conditioned reinforcer, correlated as it is with the positive and negative reinforcers discussed above. That is, most of us are quite used to evaluating, as it were, our own behavior prior to the consequation, provided by others, that may also be seen as an evaluation. Thus, if in the past Joe has been reinforced for acting in accord with some piece of verbal behavior which similarly functioned as an $S^D$ correlated with positive and/or negative reinforcement for so behaving, and the present behavior is in part a function of this history, so acting is a straightforward purpose according to (4).

The interesting case, of course, is the purpose of breaking Int's window. The pivotal aspect of this situation is the fact that the description of this purpose is entailed by the description of the above purpose, 'acting in accord with the dare to break Int's window.' That is, it is clearly impossible for Joe to act in accord with the dare without also breaking Int's window. The behaviorist cannot, of course, simply appeal to this entailment relation, saying that whenever D1 describes a purpose and D1 entails D2, D2 also describes a purpose of the
same behavior. Such an appeal is insufficient because the sort of relation involved is not expressable in simple extensional logic; that is, the entailment itself, called semantic entailment, is not obviously described without intentional language. That D1 entails D2 derives from the meanings of the non-logical terms in the expressions, not the logical terms (not, or, and, if...then, some, all). Some account of when D2 is entailed by D1 is needed.

A solution may begin by noting that semantic entailment can be explained in terms of meanings. We have already seen that Skinner can be taken to reject meaning in favor of the conditions that control the community's reinforcement of verbal responses. Perhaps, then, we can say that D1 entails D2 just in case that the sufficient conditions for the community's reinforcement of D1 are also sufficient conditions for the community's reinforcement of D2. Then, when D1 entails D2, the community cannot reinforce the use of D1 to describe a purpose without also reinforcing the same use of D2. These results can be expressed as scholia to (4):

(4a) A verbal description D1 entails another, D2, if and only if the variables controlling the community's reinforcement of D1 are sufficient conditions for the community's reinforcement of D2.

(4b) If (i) a verbal description D1 entails another, D2, and (ii) D1 describes a purpose of behavior B, then (iii) D2 also describes a purpose of B.

So far so good. Scholia (4a) and (4b) provide plausible interpretations of the relation of semantic entailment and the "impossibility" of Joe's acting in accord with the dare to break Int's window without also acting in order to break Int's window. Although this may patch
up (4) to handle this and similar cases, something is still wanting. Specifically, one might ask how (4a) and (4b), in handling these cases, do so in accord with Skinner's intuition that reinforcement takes the place of purpose. Are (4a) and (4b) *ad hoc* provisions, or can they be shown to extend the scope of (4) in an intuitively plausible way? I believe this extension is not *ad hoc*. D2 is contained in D1 in the sense, that, by (4a), the conditions sufficient for the community's reinforcement of D2 are a sub-set of those for D1. Thus, ascribing a purpose via D1 points to a reinforcement history which, in conjunction with the present situation, says something about how that history affects current behavior, and part of that something is also said by ascribing a purpose via D2. Therefore, ascribing a purpose via D2 goes no further than the ascription via D1—indeed, in general, not as far, since the variables controlling reinforcement of D2 will be a *proper* sub-set of those for D1. So the extension of (4) by (4a) and (4b) allows us to pick out additional descriptions of purpose but not, in a real sense, additional purposes *per se*.

Intentionalists will be quick to object to analyses of purpose in terms of (4), and a large proportion of these objections involve claims similar to the entailment of beliefs and/or knowledge. All such objections, of course, begin with a crucial premise about the "logic" of purposes. Roughly, all of this class of objections take the form:

(8) If it is true that \( x \) did \( y \) in order to get \( z \), then it necessarily follows that \( M \), where \( M \) is an intentional sentence about \( x, y, \) and/or \( z \).

With regard to Mo's behavior, as described in (5), a list of Ms would
include:

(a) Mo wanted a soda.
(b) Mo took the machine to be a soda machine.
(c) Mo took the coin to be a proper one.
(d) Mo knew how to operate the machine.
(e) Mo believed that the machine worked.
(f) Mo believed that his behavior would result in a soda.

The emphasized words, of course, are intentional and, would in turn be held to imply still more intentional sentences. The crucial problem involved in (8) is the notion of entailment used. It is clearly not the relation that holds between the axioms and theorems of a formal system, for whatever sentence is put in place of 'M', the denial of (8) is not a formal contradiction. Although it must be conceded that there is something distinctly odd about saying, in the typical case, that \( x \) did \( y \) to get \( z \) but \( x \) didn't believe that doing \( y \) would get him \( z \), and, moreover, he didn't want \( z \) to begin with, there is yet no contradiction in the strict sense. Perhaps this objection can be turned to advantage. For there is no convincing explanation of this oddness in the literature, and if the operant analysis can shed some light on the situation, that should certainly improve the case for the operant viewpoint on intentional language.

The "logic" between (7) and (a) is straightforwardly derived from the facts of operant conditioning. First we must translate (a) into operant language, by interpreting it as a statement about the conditions which rendered a part of Mo's history of reinforcement operative at the time he did \( y \); or, rather, a statement that some conditions obtained to make probable any of those responses in his repertoire.
which have been reinforced by \( z \) in similar situations. We are talking, of course, about such things as deprivation and aversive stimulation variables. It is surely a plausible interpretation, and provides some insight into this "logic" of purpose. For a history of reinforcement is not a sufficient condition for a response to occur. A rat, for example, will cease his bar-pressing when he has consumed a certain amount of food. If he is then deprived of food for twenty-four hours and returned to the experimental space, he will again press the bar. Similarly, even with a long history of successful behavior, Mo will not act to obtain a soda unless some causally relevant conditions exist. Two typical relevant conditions, of course, are states of deprivation (thirst or low blood sugar) and aversive stimuli (being hot and sweaty). Others might be the aversiveness provided by his comrades should he fail to get a soda for each of them, the aversiveness of having nothing to do while others drink, etc. The point is that statements like \( (4) \) are incomplete, and, within the context of the laws of operant behavior, imply some statement which is an interpretation of \( (a) \).

There is also a straightforward derivation of the "logic" between \( (7) \) and \( (d) \) from the facts of operant conditioning. A plausible interpretation of \( (d) \) is a statement that the relevant response(s) are in Mo's repertoire. He knows how to operate the machine in exactly the sense that a naive rat does not know how to operate the pellet dispenser, and a rat who has pressed the bar once and gotten a pellet thereafter does know how (c.f., Skinner, 1938). Now, if a causal factor in Mo's present behavior is a history of reinforcement for behavior of
the same class, behavior of that class is obviously in his repertoire. Thus, interpreting sentences like (3) in terms of (4) provides, within the context of the laws of operant behavior, a truly logical inference to statements which are plausible interpretations of (d).

These interpretations, of (a), (d), and the logic between them and (7), provide additional confirmation of the operant framework in several ways. First, plausible interpretations of (a) and (d) serve to meet the intentionalist challenge, thus helping to dispel doubts of the adequacy of the operant analysis. Second, by offering an interpretation of the "logic" between such sentences as (7) and (a), we find the analysis approaching a systematic nature which should give heart just as Davidson's perception of a system in philosophy's failures caused him to lose heart. And, that system is added to in a third way, because it is a fact that affirming (7) and denying (a) is odd. We assume that our linguistic practices evolved along with the constant operation of the facts of operant conditioning, and that longstanding practices in describing behavior should be in some measure a reflection of the facts. That these practices hold as odd the conjunction of what the facts reveal as impossible is a tribute both to the discriminatory sensitivity of language and to the veracity of the operant position. As we delve further into operant interpretations of linguistic patterns, we should not be surprised to find the one confirming the other.

Sentences (b), (c), (e), and (f) remain. They will require a much more complex analysis. We will examine the first two in the next section. The sentences about believing may be the most difficult,
because the term is used in so many different ways, and because philo-
sophers have, since the days of Plato, analyzed this term in so many
ways. This complex topic must be reserved for another chapter.

An Interpretation: Perception

The two sentences describing Mo's having taken a thing to be of
a given sort reveal the intentional nature of perception. Skinner's
awareness of this feature of perception is shown when he says

The heart of the behavioristic position on conscious
experience may be summed up in this way: seeing does not
imply something seen (Skinner, 1963, p. 234).

That this should be the "heart" of Skinner's position is particularly
noteworthy as we recall Dennett's observation that the task of refuting
intentionalism "...is the task of somehow getting from motion and matter
to content and purpose—and back." We have already come quite a way
with purpose; if as good a case can be made on the content side, we
will find behaviorism in a very good position indeed.

The problem of content arises when the purely physical description
of stimuli is taken seriously. To speak, as we have, of sodas, coins,
and machines begs this question, for we are bound to account for the
fact that the machine was projected onto Mo's retinas as a physical
image, so to speak, but Mo (correctly) saw it as a soda machine. He
could just as easily, with a different history (or state of deprivation
or aversive stimulation), have seen it as a martian, or, more to the
point, as a physical object that had no significance to him. Thus,
such attributions of purpose do seem to involve (typically tacit) claims
or assumptions about the agent's perception. Accordingly, while (4) might be considered an adequate account of the purposive aspect of behavior, it is considered so at the expense of utilizing intentional descriptions of the agent's perceptions when we apply (4) to particular cases, as in (6) and (7). This will not do, except as a provisional step in developing our account, for the challenge is to purge all intentional descriptions and idioms; moreover, if we are to follow Dennett's guideline, we must contrast purpose and content with matter and the motion thereof, and it is clear that we have not yet purified our interpretations to include only specifications of motion and matter.

Dennett again provides a neat characterization of the problem, and at the same time points the way to a solution. He points out that no input "...can be said to have the significance 'A' until it is 'taken' to have the significance 'A'..." on the output side.

This is not the epistemological point that as behaviorists we cannot tell whether the organism's brain has discriminated its stimulus as having the significance 'A' until the organism manifests this in its behavior, but the logical or conceptual point that it makes no sense to suppose that the discrimination of stimuli by their significance can occur solely on the afferent side of the brain...We could in principle record the afferent activity...and...determine that the animal's efferent analysis system had unique and different outputs for circles and squares...This is the sense of discrimination of interest in research in pattern recognition devices, where all that is at issue is whether or not the system is capable of producing outputs—whatever they may be—that co-occur with the critical patterns of the inputs...This is not yet discrimination by significance, however. We would not give as the conclusion of this experiment that the animal could discriminate circles as circles...Furthermore, for all animals lower than human beings there is no behavioral experiment we could perform that would have this as its conclusion, since circles and squares, even under laboratory

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
conditions, could have no bearing as circles and squares on the life and activities of the animal (Dennett, 1969, pp. 74-75, emphasis in original).

Dennett's point is evidently intended as a logical or grammatical one; hence, "makes no sense." That which he refers to as "pattern recognition" can be compared to a simplified version of discriminative respondent conditioning. For example, it has been shown that a dog can be classically conditioned to salivate when shown a circle, but not when shown a square, or *vica versa* (Kimble, 1961, pp. 361ff). If we simplify by ignoring such subtleties as respondent habituation and generalization, this models Dennett's system which is "capable of producing outputs...that co-occur with the critical patterns of the inputs" (Dennett, 1969). (One might wish to make the similarity greater by supposing a dog which salivates to one pattern and exhibits pupillary dilation to the other.) Clearly, Dennett is right in saying that this sort of discriminative behavior does not show that the stimuli have significance for the dog as squares or circles. His description by significance is of another sort. It is a commonplace that what I see in a cloud chamber image and what a physicist sees are "different". "We see the same object differently." The difference is between what Dennett calls the significance of the stimulus, and what others might call the meaning of the stimulus for the organism. Clearly, the stimuli are held constant *ex hypothesis*, so Dennett would have us conclude that the difference must lie "on the response side". But the crucial case is the one where neither the physicist nor I make any overt response at all. How can the behaviorist account for this difference in
the content of our perception?

Let us consider a simple case. Presented with the same steak cooked medium, A may respond "Too well done" and B "Too rare." Both are responding to the same physical properties, but each clearly perceived the same thing differently. Ignoring certain subtleties, this is no different from the case where pigeon A responds at a high rate to a green light, and pigeon B at a low rate. In analyzing such cases, the behaviorist expects to find a history of differential reinforcement to account for the differential responding. The stimulus has different "meanings" for different organisms because it has functioned in different contingencies in their respective histories.

The contingencies, then, determine the "bearing" Dennett refers to. No rat will respond to a circle as Dennett does because circles cannot have entered into all of the contingencies in the rat's history that they have entered into in Dennett's. Dennett has shown that the stimulus is not the critical feature; his mistake is in supposing that it must therefore be the response. Actually, we need know little of the properties of either stimulus or response, except insofar as such knowledge is needed to isolate the contingencies responsible for the operant in question. There must be a stimulus to be taken, and some response, but the response indicates that the organism took the stimulus as thus and so only insofar as it indicates which contingencies led to that particular piece of behavior.

For example, if bird A responds in a "scalloped" manner, peaking every minute, we can conclude rather safely that the stimulus signifies
the prevalence of an FI 1 minute schedule. Note that the specification of purpose or reinforcer is not essential here, which verifies the dual (purpose and content) nature of intentionality as first enunciated by Brentano, and explicated by Dennett.

Returning to the cloud chamber, suppose neither I nor the physicist respond overtly. All that is needed is some covert response—again Dennett is right that some response is needed, but only to mark the taking of the stimulus in some way or other—which is (a) controlled by the chamber image as discriminative stimulus, and (b) due to different contingencies in his case than in mine. In principle, at least, the responses could be executed with the same part of the body—say, neuron number 351 in some neurological lexicon. The difference lies in the contingencies responsible for the operant, rather than either the stimulus per se or the response topography. Thus, a rat will respond to a circle as a circle just as long as his response to the figure is due to the same contingencies as Dennett's so often is (a very unlikely eventuality, to be sure).

The present analysis diverges from our ordinary approach in which we assess whether or not I took the cloud chamber image as a physicist would by the response(s) I can make. This is not usually wrong, in that it does not often lead to mistakes, since response forms in our culture usually correlate with contingencies rather well at this level. That is, our culture's contingencies typically involve the specification of response topography. Further, we usually can respond to some covert response of our own, and thus be "aware" of how we take a stimulus
(although more needs to be said about awareness). This is also rarely mistaken, presumably because contingencies shape different forms for such taking responses, and other contingencies shape rather tight discriminative repertoires of naming such takings. Still, the extreme cases show that the crucial distinction lies in the contingencies themselves, rather than what they shape.

**Summary**

Beginning with Dennett's analysis of the problem of intentionality in terms of a problem of getting back and forth between matter and motion on the one hand and purpose and content on the other, we have offered interpretations of purpose and content in terms of causally relevant histories of reinforcement and contingencies. In the process, we have found a system of success to contrast with Davidson's system of failures, in as much as certain "logical" qualities about intentional descriptions have corresponding qualities on the side of matter and motion. This system was provided by interpreting other intentional phenomena than purpose and content, such as that described by 'x wants y.' Two critical phenomena remain: beliefs and awareness. They are dealt with in the following chapter.
CHAPTER IV

Awareness and Belief; A Beginning

Several intention idioms were interpreted in the last chapter. However, it must be recalled that the force of that chapter rests to a great extent upon Dennett's analysis in terms of purpose and content. Many philosophers may find fault with his argument, and will then demand to know why belief and awareness are not regarded as just as important or more so. Whatever the status of Dennett's view, it cannot be denied that belief and awareness provide challenging and widely cited examples of the intentional. We may find that we are unable to offer conclusive interpretations here, and that we must be content with a beginning.

Awareness

The topic of awareness or consciousness has been discussed so often by so many that we must begin by distinguishing various senses, and specifying the sense to be discussed here. Several senses are illustrated by the following:

(a) I am aware of your concept of love.
(b) Yes, now that you mention it, I am aware of your discomfort.
(c) I was aware of your discomfort—it clearly seems to have been the cause of my solicitousness—but I don't recall being aware of it.

Sentence (a) is not at issue here; this sense of awareness is something like 'I know what to say about love to speak as you would.' Sentence

45
(b) is closer; however, insofar as it marks my attention to a stimulus, it is not to the point either. As Skinner (1953) has noted, (b) merely marks my behavior's having come under stimulus control of the stimuli which I take to indicate your discomfort. That this can occur without awareness in the present sense is shown in (c), a seeming contradiction. The equivocation on 'aware' is revealed by the qualifying phrases, however: I certainly was aware in that my behavior was controlled, but I was not aware (granting for the moment the veracity of my recall) in that something else that might have occurred did not. It is this "something else" that is at issue.

This "something else" can be illustrated by comparing:

(a) I saw the punch coming, and responded by blocking it.
(b) I didn't really see that punch until after I'd blocked it.

In both cases, visual awareness of the attention sort is indicated by the block, presumed to have been controlled by the punch as stimulus. But the "something else" is missing only in (b). Another case (which I owe to Dr. Mountjoy) is found in the unconditioned withdrawal reflex elicited by the heat of touching a hot pan, where we often say we are not aware of the pain until a discrimination interval after contact and withdrawal. To be sure, the unconditioned nature of withdrawal makes this case somewhat primitive, but it is therefore graphic: attention a la operant stimulus control is absent, and the time lag proves that there is "something else," something beyond a more general scope of stimulus control, even in this non-visual mode. It is this "something else" which is the awareness we shall examine in this section, and we shall understand 'awareness' and other similar idioms in this sense.
for the rest of this section.

One problem is the difficulty of describing the details of awareness. We cannot even do this by contrast, since it is so ubiquitous. And we cannot look at it somehow, as it seems to be one mode, at least, of looking at or examining closely. The visual mode certainly suggests images, a sort of movie we watch in addition to responding; and, to continue to metaphor, the auditory mode simply suggests a similarly redundant soundtrack. The metaphor receives further validation of a sort in the "feelies" of Huxley's *Brave New World*, where movies are imagined that also stimulate tactily and olfactorily, and the idea is easily extended to such inner sensations as intestinal activity produces. Unfortunately, however, much of this sort of metaphor satisfies those who find awareness an intuitively clear notion, it helps little for those who do not.

Another problem is suggested by our label "something else." Whatever role awareness plays in our repertoire, it is clearly not necessary in many cases, and seems unnecessary in many, many more. Just as one can see a punch coming without being aware of it, it seems likely that one could read this paper without being aware of seeing the visual stimuli (and without being aware of reading it). Yet awareness is so nearly ubiquitous. Any operant analysis must ultimately appeal to some payoff, some reinforcement for awareness, but the most obvious candidate—the social reinforcement for reporting the contents of one's perception—seems a rather paltry payoff for such ubiquitousness. How can an operant analysis interpret a behavior—surely it is a behavior—that seems
to deny the laws of operant behavior?

One approach, which I take to be Skinner's (c.f., Skinner, 1963a), holds that awareness is a response, a discriminative response, to the response or other event of which the organism is said to be aware. For example, our model of perception has it that vision begins with a series of reflexes of the respondent type. Sooner or later, a respondent response will be or produce a discriminative stimulus for an operant response, which may be labelled the seeing response; the respondents may be labelled copies or sensation. Skinner suggests that awareness is seeing that we are seeing, or seeing what we are seeing, which suggests that the seeing response in turn is or produces a stimulus controlling another operant response, the response of being aware of what is seen, or the awareness response. Of course, the seeing response will also be a stimulus for other operants, ending in some overt behavior in many cases.

Skinner seems to hold that the awareness response is a social one, shaped and maintained because the community is sometimes reinforced by how it seems to us, etc.

"Knowing" or "being aware of" what is happening in turning a handspring involves discriminative responses, such as naming or describing, which arise from contingencies necessarily arranged by a verbal community (Skinner, 1963a, p. 229).

But it seems to me that my awareness of the pattern in my living room carpet does not allow me to report how it seems to me. In this case, I am hard pressed to but admit that I am aware of the pattern, but I cannot say much of anything about it. To be sure, I can report that I see it, that it is very complex, etc., but the point is that there
seem to me to be aspects of that of which I am aware which I cannot describe.

A puzzling aspect of awareness is that it is so ubiquitous. I am almost never not aware of some part(s) of my environment. To be sure, intermittent reinforcement can account for very persistent behaviors, but I am aware even where there is nobody to report to. Perhaps awareness has some advantage in that it enables me, so to speak, to report to myself, thereby bringing the private events into the same kind of play that the public ones presented by the external environment and the verbal community in particular impress upon me.

One may be tempted, in this puzzlement, to deny that awareness as an additional response is needed. However, we do occasionally respond without awareness, as when I avoid something and say I wasn't aware of that which I avoided until after I had already avoided it. We sometimes label such behavior reflexive, thereby denoting the absence of awareness. It seems, then, to be something extra indeed.

Skinner's frequent objection to copies (e.g., 1963a) seems to be in error. To be sure, his major point, that one must eventually see, wherever the copy is placed, is unobjectionable. Nonetheless, it is clear that I can recall a scene, seeing it again and being aware thereof, and read off the details of the scene. Thus, it must be that the details are represented in the seeing response, which demands that the details be represented in the response we have called sensation.

Indeed, it is the apparent representation of detail in awareness that is hard to account for. Why not report to yourself, even though
I do not make the awareness response; that is, why not be aware only when needed, and then be aware overtly? Can we simply say that the awareness response feels different than describing our environment because it is not overt, or because of its locus in the nervous system? These problems remain as we leave this topic, but they are problems, not refutations of behaviorism. This much can be said in view of the systematic success of the previous chapter.

**Beliefs**

Philosophers have had, since the time of Plato, an inordinate interest in beliefs, what they are, how we come by them, how they relate to action, and so on. The idea that there may not be any such thing as a belief has resisted until relatively recent times, when philosophers began to explore a behavioral approach. Beliefs have been interpreted as tendencies to behave in various ways (e.g., Sellars and Chisolm, 1958), but, as with perception and awareness, the presumed fact that we can be said to believe p even though we exhibit no correlative specific behavior has held against such analyses.

Perhaps the first step towards clarifying a behavioral analysis of beliefs is to distinguish three types on the basis of the variables which account for their ascription to another or to oneself. Two of these are described by Skinner (1957) in terms of response strength and autoclitic variables. A third possibility is a strictly intraverbal dynamic. The first two are easily confused, and the third is often mistakenly subsumed under the first, which is the type that should
truly concern philosophers interested in epistemology.

Let us begin with the intraverbal sort. Intraverbal behavior is defined (Skinner, 1957) as verbal behavior, under the control of other verbal behavior, where there is no point-to-point correspondence. Thus, I may say that I believe this or that when asked why I want to do thus and so, because past experience has shown that such a response to such a stimulus will be positively or negatively reinforcing, for example, by making me appear wise or freeing me from threats of appearing stupid. Intraverbal beliefs are not truly beliefs in the ordinary sense, although I may report true beliefs in this sort of context. But when my proclamation is not under the control of my own behavior, or the strength thereof, it is purely intraverbal, although there may be some mand-like qualities involved in getting another to behave in certain ways. The major point is that there are no tact-like qualities involved, even though they are often mistakenly attributed.

The autoclitic kind of belief is reported under a tact-like stimulus control of the strength of my own behavior. This is to be contrasted with such utterances as 'I know that p' and compared with such as 'I think that p.' The autoclitic 'believe that' is controlled by the strength of my response (p). Insofar as there is a correlation between the strength of this verbal response and the strength of nonverbal or other verbal responses that may be called 'acting as if p were true,' the autoclitic provides evidence of my epistemic belief(s); however, many autoclitic beliefs concern propositions bearing little or no relation to most of my behavior.
The "true belief" as I have called it or the epistemic belief, is ascribed by another to me on the basis of my behavior, often my non-verbal behavior. To explicate this notion, we must turn to Skinner's distinction between rule-governed and contingency-shaped behavior (Skinner, 1969, 1974). The key here is an interpretation of what 'acting in accord with rule R' means. To begin with, a rule is a tact, controlled by, or "describing" a stimulus situation and a response, and sometimes a consequence; thus, a rule is a tact under the control of a contingency. One is said to act in accord with a rule R to the extent that if (i) one is motivated to obtain/avoid the relevant consequence, then (ii) one makes the response tacted in the stimulus situation tacted in R. For example, 'waste not want not' tacts a very general situation and a very general consequence; I behave in accord with it just in case wasting is a current aversive stimulus for me and I have the opportunity to waste not.

Now, I may behave in accord with R either because my behavior has been shaped by the contingency tacted by R, or because R is a stimulus for my behavior. That is, I may act in accord with R either under the control of the rule as a verbal stimulus, or under direct control of the contingencies themselves. It is a purely philosophical question whether epistemological belief involves one, the other, or both of these types of behaving in accord with p, where p is a tact of a contingency. But it is clear that the empirical question involves both: we may attribute a belief on the basis of the behavior, even if we do not know whether the behavior is rule-governed or contingency-shaped.
It should be noted that this interpretation of 'acting in accord with R' is a quantitative one; that is, one acts in accord with the rule to some extent or other. Specifically, the strength of my belief is proportional to the strength of my tendency to respond in accord with R under the relevant conditions, which makes it proportional to the probability of appropriate behavior. For some time, philosophers have tried to interpret beliefs in terms of the probability calculus, and this may be a relevant entering wedge. In any case, it does seem to be yet another connection between the intuitions of philosophers and behavioral psychologists, and should be yet another point to consider in dealing with Davidson's view that there has been a system in the failures to provide behavioral interpretations. Perhaps it may also provide some validation of Skinner's frequent but unsystematic use of probability conceptions for his description of response strength. (Incidently, this view also correlates with Schlick's (1934) view of scientific confirmation, and, thus, leads to a formulation of 'true belief' in the epistemic sense as a rule which, if consistently followed, would lead consistently to reinforcement.)

By way of objection, one may inquire as to the nature of the relation(s) among the rule, the contingency, and the behavior. Specifically, how does one know that a rule describes this rather than that contingency, or that this situation does or doesn't fit the rule, or that this or that response or consequence fits the rule? This is, perhaps, most urgently asked when one recalls the possibility of multiple descriptions of physical and/or behavioral events.
which are all appropriate in the sense that the community would rein-
force them. The key here is the dimensions which will lead to the 
community's reinforcement; for example, a response or a stimulus will 
fit R just in case the response or stimulus parameters would lead 
the community to reinforce their description with the terms of 
R.

Another objection concerns the case where my behavior can be said 
to fit many rules. For example, the rat's bar-press may signify his belief that, by so behaving, he will obtain food, or that he will avoid punishment. Obviously, this case shows that the causes of the response must involve either the contingencies themselves or a rule, and that such causal antecedents are relevant. Thus, we see that the agent's intentions figure into the description of his behavior, and the content of his perception is also relevant. For if I mistake the stimulus situation as falling under R when it really falls under a different rule R', my behavior should be taken as evidence of my belief in R, even though my behavior may seem to also fall under R'. In sum, when the content of my perception and the purpose of my behavior is taken into account, my behavior can only be taken as evidence of one of any set of contrary beliefs. Regarding alternative rules that are not con-
trary to one another, the situation is similar. Thus, the baseball player's behavior in catching a long fly ball may be taken as evidence of his belief in a system of mechanics based upon the fact that objects near the earth fall at a rate of 16 feet per second squared, or one based on a more general law of gravity. Actually, of course, in the typical
case either choice would seem absurd. It would be absurd because the typical player has had no experience with either system and, if he had, that experience plays no part in his fly catching behavior. Again, upon referring to the organism's history, the specific relevant variables and history will point rather narrowly to only one proper attribution of belief. In this particular case, the relevant history is likely to involve previous fly balls only, and is likely to lead to some such rule as 'with balls that look like that, I can catch them by behaving in this way.'

Philosophers will want to know how to do two sorts of things: in general, what beliefs are indicated by observed behaviors, and what behavior(s) do we seek to verify 'x believes that p'. If we assume that the relevant sense of the term involves behaviors controlled by verbal stimuli, somewhat satisfying answers are possible. Observed behaviors, of course, indicate no belief of this sort unless they are controlled by some verbal stimulus R, and/or they arose from some such R. (This last qualification allows as beliefs those indicated by behaviors not presently controlled by R, but which would not occur without historic control by R. For example, I was taught to type with at least two layers of paper, lest the roller be damaged, but I currently add a "backing sheet" under the control of those stimuli controlling typing, not under the control of any R I am currently exposed to, internally or externally.) Thus, the controlling variables must be known, or inferred. Given, then, that observed behavior B arose or was emitted under the control of R, where R is specified in detail by its occurrence(s)
as a stimulus taken by x, a belief in R is indicated by B.

Now, suppose it is asked whether or not x believes that p, where p is a particular belief "expressable" by a particular verbal stimulus R, or any of several stimuli belonging to a class of stimuli of which we say "they all express the same proposition," which means simply that they are under the same control of the tact variety, although audience variables, for example, may differ (we do not intend, by so defining analyticity, to be held to any thesis regarding the existence of truly analytic sentences). The interesting case here is where x exhibits no overt behavior under the control, present or historical, of R, but x still may believe that p.

One tack is to reply that the case is absurd, and, indeed, it is hard to imagine believing that p having absolutely no effect on x's overt repertoire. The question is empirical, provided that an independent account of 'x believes that p' is provided. To provide this, three tacks remain. The first is to carefully peruse x's entire history, to ascertain whether or not there were any events which, given specifiable state of deprivation and/or aversive stimuluation, would lead to relevant overt behavior. The second is to insist that there be some covert behavior controlled by R, such as saying "R" silently to oneself. Finally, noting that 'x believes that p' describes a state of x caused by some exposure to contingencies relevant to p, 'p,' and/or 'x believes that p' could be confirmed by examining the relevant state(s) of x. There must be such a state, for it surely controls x's finally saying 'I believe that p.'
It is noteworthy that tacks one and three are barely distinguishable, since the events sought in one are just those responsible for the state in three. Tack two might be rejected, were it not for Dennett's argument, which implies that some response was necessary in x's taking his experiences in such a way as to lead to the disposition to behave indicating belief a la tacks one and three. Still, it might be possible to never have behaved under the control of R, even though one is currently in a state disposing one to do so, given only the motivation. Thus, armed with one and three, we can at least imagine x believing that p without ever having acted on that belief, and the question can be left as an empirical one whether or not this imagination is ever realized.

Summary

The traditional concepts of awareness and belief were examined. Awareness was found to present certain problems that remain unsolved. The nature of its detail and ubiquitousness are hard to account for in particular. The notion of belief was divided into three types, and only one was found to be relevant to the traditional philosophical enquiry. This was interpreted as a description of the probability of a response from a given class, given a member of a particular class of stimuli, and relevant deprivation and/or aversive stimulation. This account was found to correspond well with many traditional accounts.
CHAPTER V

Ontology

Most behaviorists espouse some sort of materialist ontology, explicitly or implicitly, and they generally take their success in empirical work as support of this view. Philosophers, on the other hand, are wont to argue that behavioral materialism is an assumption of the empirical work, and cannot be confirmed by it. Hull, in an exceptionally clear paper (Hull, 1973), presents an argument which allows us to see the relation between intentionalism and the view that ontology precedes causality.

Hull's Argument

Hull's argument is addressed to what may be called a meta-metaphysical question, namely, what sorts of facts are relevant to the metaphysical questions referred to by philosophers as "the mind-body problem"? What would count in favor of the view, for example, that there are non-spatial or in some other way non-physical minds? Or in favor of some alternative view? Hull begins by defining the most prevalent philosophical views on mind, and argues that scientific facts about causal relations are all but totally unrelated to them. His characterizations of these views are illuminating, for they enable one to see with clarity the thrust of intentionalism, and its importance in the formulation of radical behaviorism; and, incidentally, it shows how Skinner's discovery of the operant relationships figures so crucially in an understanding of intentionalism and
intentional phenomena.

Hull's terms may be unfamiliar to psychologists. An *individual* is defined "...in a broad enough sense to encompass particulars, substances, events, states, or processes..." and *characters* are defined as "...properties and relations..."; individuals are to be identified as "mental or physical" according to whether they instatiate or exemplify mental or physical characters, respectively (Hull, 1973, p. 67). Hull presumably means some such thing as 'of the realm of physics' by 'physical,' although he does not make this explicit. His use of 'mental' however, is clear enough for the behaviorist. He holds that mental and physical characters "...are detected in different ways...", suggesting different dimensions for each, and thus rendering the mental "theoretical" in a way to which Skinner has objected (Skinner, 1950). This suggestion is confirmed when Hull says that a state of consciousness ("...e.g., an instance of feeling pain, having an after image, thinking about a problem...") does not "...have location in space..." (Hull, 1973, p. 71). And again, "...any process in the brain is at a place where no mental event can be found..." (Hull, 1973, p. 71), presumably because "mental events" cannot be found at any place at all. In short, Hull's mental individuals are just the sort of thing the existence of which is denied by radical behaviorism.

Having defined his terms, Hull uses them to define five positions quite clearly:

**H1.** *Eliminative materialism:* All characters are physical; there are no mental characters; there are some individuals which exemplify physical characters.*

**H2.** *Reduction materialism:* All characters are physical;
there are some mental characters; there are some
individuals which exemplify mental characters, and some
individuals which exemplify physical ones."

H3. "Double-aspect materialism: Some characters are physical;
some characters are mental and are not identical with
any physical characters; there are individuals, all of
which have physical characters and some of which have
mental characters that are not identical with any physi­
cal characters."

H4. "Dualism: Some characters are physical; some characters
are mental and are not identical with any physical
characters; there are at least two individuals one of
which has all physical characters and one of which has
all mental characters; some of the latter are not
identical with the former, and vice versa."

H5. "Psychophysical Correlation: For every mental character
there exists a physical character such that, for any time,
some individual has that mental character at that time
if and only if some individual has that physical character
at that same time." (H1-H5 from Hull, 1973, pp. 66-67)

Since H1 alone denies the existence of mental characters, and
thus mental individuals, it is the only one of the five that a radical
behaviorist could accept. The interpretations of Chapter III may
seem to make H2 acceptable, but it must be remembered that 'mental'
is reserved for characters that are non-physical, or exemplified by
individuals which are non-physical, and the very act of interpretation,
although it allows and even relies on the use of terms traditionally
taken to denote the mental, assumes that there are no non-physical
characters or individuals. And H3-H5 are not acceptable to the
radical behaviorist for the same reason.

However, allegiance to H1 commits one to a decidedly unusual posi­
tion regarding psychophysical correlation (PC),—as expressed in H5.
For, as Hull points out, if H1 is true, then H5 must be true, but
only vacuously so. So, H1 denies that there are any mental characters
to correlate with physical ones. Thus, H1 entails that no evidence for or against PC is possible. H5 remains true because H1 entails the falsity of the antecedent of the conditional statement expressing H5's logical form (If there is mental character, then there is a physical character such that...). And, by convention, conditionals with false antecedents are always rendered true in formal logic. So, Skinner's presumed commitment to H1 deprives him of any empirical grounds for holding PC.

I believe it is clear that Skinner does subscribe to PC in a non-vacuous way. Moreover, I take it that Skinner believes that his radical behaviorism entails PC and that, in virtue of this entailment, evidence for PC provides further confirmation of his behaviorism. Indeed, PC is, for Skinner, near the heart of behaviorism. Repeating an argument against methodological behaviorism which appeared as early as 1945, (Skinner, 1945), he says in 1974

A science of behavior must consider the place of private stimuli as physical things, and in doing so it provides an alternative account of mental life. The question, then is this: What is inside the skin, and how do we know about it? The answer is, I believe, the heart of radical behaviorism (Skinner, 1974, p. 212).

This heart, of course, is to reject methodological behaviorism, which admits mentalism is beyond science and thus admits the mental, and to provide an "alternative account of mental life" in terms of behavior. The account Skinner provides has it that everything "...a person comes to know about himself..." via introspection "...is just more stimuli and responses" (Skinner, 1974, p. 216), and that "Thinking is behaving. The mistake is allocating the behavior to the mind" (Skinner, 1974, p. 104).
Thus, what we identify or label as thinking, feeling, and so on are behavioral variables. And it is plain that there must be, therefore, a correlation between "mental events" such as thinking and physical behavioral events such as stimuli and responses; for, in Skinner's view, the behavioral and the mental terms name the same events. If this were not true, indeed, the whole business of interpretation would be senseless.

Not only is Skinner committed to PC, as argued above, but it would be misleading to overlook his most explicit avowal of PC. Again while objecting to methodological behaviorism, he asks "What about psychophysics and the mathematical relation between the magnitude of stimuli and sensations?" (Skinner, 1974, p. 15); here, he rejects methodological behaviorism for ignoring psychophysics, and notes that his radical behaviorism makes a place for it. Specifically, "...private events...are the collateral products of environmental causes..." (Skinner, 1974, p. 220), and a "...psychophysical function...represents facts about the discriminative control of stimuli" (Skinner, 1974, p. 78). This cannot be viewed as any but an explicit affirmation of the research tradition associated with PC.

Clearly, then, Skinner does subscribe to a non-vacuous version of PC, and he views PC as an integral part of his position, or entailed by it. Another part of his general position concerns the relation between a science of behavior and physiology. We have no need here to discuss his views on this matter, except to note the relevance of PC. He holds that the view that psychology must eventually be absorbed
into physiology overlooks "...the possibility of a behavioral science and of what it has to say about feelings and introspectively observed states" (Skinner, 1974, p. 214). In discussing dreams, he argues that the correlations between dreams and rapid eye movements "...seem to confirm..." his interpretation (Skinner, 1974, p. 85). Perhaps the best summation of his view comes under his heading "The invasion of privacy":

Some recently developed physiological techniques seem to testify to the reality of private experiences. They seem to provide independent evidence of what introspective psychologists have been talking about. Subjective experience has become, so to speak, an objective fact. But the evidence is important for a different reason...The new evidence simply points to the fact that what is experienced introspectively is a physical condition of the body, as a behavioristic theory of knowledge has always contended (Skinner, 1969, p. 262).

The more evidence produced to confirm PC, the more Skinner's view is confirmed. It would seem, then, that something is amiss; perhaps Hull's definitions are miscast, or perhaps he simply failed to define the position of the radical behaviorist. Actually, Hull is remiss in both directions, and for an understandable if not totally defensible reason. If this reason can be found and the mistakes corrected, we may be able to avoid Hull's conclusion, namely, that Skinner's position on the mind-body problem, H1, is logically incompatible with, and thus hardly confirmed by, psychophysical correlations.

**Hull's Bias Removed**

The fault lies with one of Hull's presuppositions, and it is this axiom that leads to his biased, if not faulty, definitions. Specifically, the problem is his blind acceptance of intentionalism. He refers
...the arguments of philosophers which show that there are non-causal, non-behavioral, non-linguistic features of cognitive acts which cannot be redescribed in physical terms—that is, the intentional features of cognition. (Hull, 1975, p. 65)

The arguments he cites, however, "show" no such thing; they merely point to certain difficulties that philosophers have had in physically redescribing these intentional features (Sellars and Chisolm, 1958).

It is not surprising that an intentionalist would overlook the possibility of the following reformulation of H1:

RH1. Radical Behaviorism: All characters are physical; some characters are intentional; there are no mental characters; there are some individuals which exemplify physical characters; there are some individuals which exemplify intentional characters.

There are two correlaries of RH1:

RH1a. Every intentional character is identical with some physical character.

RH1b. Any individual which has one or more intentional characters is identical with an individual all of whose characters are physical.

Similarly, the intentionalist might be expected to overlook the following reformulation of H5:

RH5. Psychophysical Correlation: For every intentional character there exists a physical character such that, for any time, some thing has that intentional character if and only if some thing has that physical character at that time.

The intentionalist will see little to choose between in H1 and RH1, or H5 and RH5, for he will take intentional characters to be as ontologically esoteric as (or even identical with) mental ones. But RH1 and RH5 put the issues in sharper focus. The former entails the
latter, so that evidence for it is *ipso facto* evidence for radical behaviorism, just as it should be. Moreover, the correlative RHla clearly shows that the crux of the matter is intentionalism, for if it can be shown that intentional language expresses nothing more than the physical language of behaviorism, the identity of the two kinds of characters is fixed. Finally, radical behaviorism is seen in a rare light, countenancing intentional expressions, but well within the framework of an eliminative materialism.

Of course, we have already seen radical behaviorism in this light, and we have already seen the foundation, at least, of a proof for RHla'. Every intentional character is identical with some operant character.

where we understand by 'operant character' just those properties and relations unique to the operant analysis of behavior. All that remains, then, is some demonstration that

RHla". Every operant character is identical with some physical character.

is true. Oddly enough, there are those who seriously doubt the truth of RHla", and we shall turn to a discussion of this in the next chapter. For now, it is enough to note that, by reformulating Hull's definitions, we have found a defense for the view, common to many behaviorists, that evidence for psychophysical correlations is evidence for a behavioral formulation of eliminative materialism.

**Summary**

Hull's argument, based upon his definition, was that the causal
evidence of the sort that bears on the thesis of psychophysical correlation does not bear on the metaphysical theses on the mind-body problem. These definitions were seen to be biased by Hull's assumption of the truth of intentionalism, and revised definitions were proposed along the lines of radical behaviorism as interpreted in Chapters II and III. These new definitions were seen to change the picture significantly, such that causal evidence for the thesis of psychophysical correlation would bear upon metaphysical theses on the mind-body problem. In particular, the truth of the thesis of psychophysical correlation, as revised, was seen to be a logical consequence of the definition of radical behaviorism. The definition of radical behaviorism was also seen to entail the identity of intentional characters with operant ones, as well as the identity of operant characters with physical ones.
CHAPTER VI

Operant Terminology

The purpose of this chapter will be two-fold: to provide some demonstration of the physical nature of operant variables, thus providing some demonstration of the truth of RH1b from the last chapter; and, in so doing, to provide a missing link in the argument of Chapter III. In that chapter, we examined Dennett's notion that the problem of intentionality can be solved by getting from matter and motion to content and purpose, and back again. The argument, thus far, has only provided interpretations back and forth between operant descriptions and those of content and purpose. To provide the link with matter and motion, we must define operant terms by reference only to physical properties and relations. The upshot, then, will be to strengthen the relation between the arguments of Chapters III and IV and, thus, the relation between the problem of intentionality and radical behaviorism's materialism.

Basics

In 1931, Skinner surveyed the uses of the term 'reflex' from the time of Descartes, and analyzed the common thread throughout these uses as a simple correlation between stimulus and response (Skinner, 1931). In 1935, he dealt with the problem of specifying the properties which empirically define these terms (Skinner, 1935a). Because stimuli and responses varied from one correlation to the next, he offered a refinement of the earlier definition: a reflex was a correlation between classes of stimuli and classes of responses, a class of an ordered

67
pair of classes. Complete generality had to be avoided, of course, unless the scientific analysis of behavior into parts was abandoned. The question of how general the classes should be made was answered by Skinner in an empirical manner: define the smallest classes that lead to orderly changes in the correlation.

This appeal to second order laws—the correlation itself is nothing more than a first order, descriptive law—is found in the third stage of Skinner's development of the concept of the reflex, in which he distinguishes operant reflexes from respondent ones (Skinner, 1935b, and 1937). The basic distinction was made in terms of the order of the pairs: the key respondent stimulus precedes the response, the operant follows. There is some debate whether this way of making the distinction will do, for the prior response in the operant case must occur in the presence of some prior stimulus. Some have argued from this fact that it is logically impossible to arrange an operant R-S correlation without setting up the conditions for conditioning an S-R respondent reflex (Hilgard and Bower, 1966). And it is true that the establishment of an operant correlation almost inevitably creates a class of three ordered classes: an S-R-S relationship. Still, the distinction can be made in terms of second order laws, which are very different for the two types of reflex. For example, intermittent "reinforcement" of a conditioned respondent does not result in the marked increase in the time required for extinction (the removal of all reinforcement for a given response class) to reduce the response to its original level that intermittent operant reinforcement does.

It is assumed that the reader is familiar with the basic phenomena
of operant and respondent conditioning (for a review, see Skinner, 1953). Nonetheless, a few aspects of the operant concept deserve mention. For example, using the respondent as a model, with its typically narrow response classes, one is likely to expect similarly narrow operant classes. But the classes that constitute an operant may be quite large. For example, in learning to drive a car, the novice may emit several operant responses under separate stimulus control—depress clutch, let up on accelerator, grasp shift lever, etc.—but the accomplished driver may emit only one: shift to second. Differences between the two performances could be shown empirically by finding an essential stimulus for the novice's developing behavior the removal of which upsets his performance—say, the roar of the engine that results from not letting up on the accelerator—and finding that the absence of this \( S^D \) does not affect the accomplished performance. Other methods are available for the specification of large response classes, such as having the subject begin in medias res; for, although chains can be begun in the middle without upsetting behavior, behavior units presumably cannot (virtually by the definition of a unit). Verbal behavior provides clear enough examples here, such as the person who can recite a long poem correctly, but only by starting at the beginning.

Similarly, stimulus classes—discriminative and reinforcing—can be quite large. At the unconditioned level, 'food' specifies a broad class of stimuli that reinforce the behavior of the very hungry person (i.e., one who is greatly deprived of food). Conditioned stimuli are also often quite broad. A painting may function as an \( S^D \) for the verbal
response "Ah, a French Impressionist," or reinforce a lover of the Ashcan school for his diligent search of the museum for the proper room. The separate "points" of a masterpiece of Pointillism may have been separate stimuli for the painter, although even that is questionable, but it is the distribution and colors of these points on a canvas that reinforce the appreciative viewer, not the separate points one-by-one. Indeed, the CS of a respondent may be quite large; Kimble (1961, p. 207) calls Pavlov's earliest work "crude," by which he evidently means that the CS was not specified any more narrowly than "showing the dog a piece of bread."

This possibility of broad classes should not be taken to countenance the uncritical use of terms from ordinary language; Skinner realized this as soon as he went to the class definition (Skinner, 1935a). Ultimately, events must be specified in terms of their physical dimensions, which it is clearly impossible to always accomplish at this time. Thus, ordinary talk about events is relied upon so that the enterprise can get off the ground, but that language is restricted as much as possible to terms which have a high degree of inter-rater reliability. Intentional language is explicitly barred, except in those cases where the direction of a translation into physical language is quite clear. Indeed, even the use of ordinary physical language is restricted to the most reliable usages in empirical research and application. The problem is that, especially in the application of operant technology, strict adherence to physical terminology renders impossible the attempt to work on certain problems that demand work. A general rule of thumb is to talk in as nearly a pure physical vocabulary as
possible, but talk. If this looseness is in fact unjustified in any particular instance, the data will eventually show it in their surprising character. As shall be seen below, there are certain "theoretical" contexts in which this rule is relaxed to a large extent, and with good justification.

Some Niceties

The definition in terms of classes of classes is often ignored. One often speaks of an "operant response," which is ambiguous as to class versus member, or an "operant response class," or even an "operant." However, an operant response class is defined in terms of the properties upon which the reinforcement of its members is contingent. Therefore, its defining properties, implicitly or explicitly, include a specification of a topography and its relation(s) to time, prior stimuli, and all of the variables that define a particular reinforcement schedule. For example, 'bar press' does not adequately define an operant, except in a context where the presses' topography (e.g., force, which paw, inter-response time, duration of depression, etc.) and the schedule under which they are reinforced (e.g., a multiple FI 1' extinction schedule for 3" access to grain, where the FI leg is accompanied by a 1,000 Hz tone of 60 db, and the grain presentation by a loud relay and a light bulb) are clearly specified. Indeed, the organism's history is relevant as far back in time as events can be found that are functionally related to topography or response rate.

Another common misunderstanding stems from the use of 'stimulus'
and 'response,' which suggest a dichotomy between organism and environment. The fact that operant studies are usually executed in a way that may seem to tacitly accept this dichotomy probably adds to this confusion. And Skinner's opposition to "theories" (Skinner, 1950) and "inner states or agents" (Skinner, 1953, p. 35) is also evidently relevant. But stimuli and responses are defined by the contingencies which generate and maintain them, not by their spatial loci. Thus, Skinner has held that the "skin is not that important as a boundary" (Skinner, 1963a, p. 228), and began talking rather freely about private events at the same time he first spoke out against the use of Bridgeman's operationism regarding psychological terms (Skinner, 1945). Similar formulations of private events appeared in 1953, 1957, and 1969, to cite a few.

This confusion can be cleared up best, perhaps, by clarifying the objection to "theories of learning." Any interpretation is open to debate here, but the one offered is not without merit. The objection is not to "theories" in the sense of explanations in terms of unobservables; rather, one might say it is to explanatory terms naming the wrong kind of unobservables. Specifically, "...any explanation of an observed fact which appeals to events...described in different terms, and measured, if at all, in different dimensions..." (Skinner, 1950, p. 39) is objectionable. Compare his position on private events:

Science often talks about things it cannot see or measure. When a man tosses a penny into the air, it must be assumed that he tosses the earth beneath him downwards. It is quite out of the question to see or measure the effect on the earth, but the effect must be assumed for the sake of a consistent account. An adequate science of behavior must consider events
taking place within the skin of the organism, but as physiological mediators of behavior, but as part of behavior itself (Skinner, 1963a, p. 228, emphasis added).

The functional relation between "action and reaction" is a well understood one independent of the penny toss case; traditional theories of learning are not. Rather, they explain behavior in terms such that (i) they are not observable, and (ii) events of the same kind have never been observed to have the kinds of effects which constitute the explicandum. The functional relations among stimuli and responses are understood in some detail, so they may be taken to occur when and where necessary for the sake of a consistent account. Such assumed parts of behavior remain assumed, and thus in an acceptable sense theoretical. They haven't the same status as actually observed events; they partake of a kind of speculativeness. But they are not completely ad hoc, since our knowledge of the kind of event in question prevents us from assuming properties not observed elsewhere.

This presumed irrelevance of the skin as a boundary, as well as the variability of the size of classes allowed by the definitions, has been vindicated by empirical work. Perhaps the smallest conceivable response class—firings of single neurons—have been apparently conditioned several times (Olds, 1965; Fetz, 1969). Operant conditioning has also been shown to be effective in generating a response class that properly deserves the name 'creativity.' Porpoises were reinforced if and only if they exhibited a bodily movement never before observed by the trainers, and they responded in a manner that left the establishment of this class practically unquestionable (Pryor, 1969). In terms
of topographical variance, it would be hard to imagine many larger classes.

Ontology Again

It may seem that the concept of the two reflexes is difficult to use in interpreting certain empirical practices among operant conditioners, and in this way the empirical work may be said to undermine the definitions. Before discussing some of these problems, it will prove useful to clear away some philosophical brush. It will be held, as is typical in science, that observable relations such as reflexes are "real" in the sense that tables and chairs are real. Such logical theories of perception as phenomenalism are eschewed, for a science of psychology must provide a causal account of perception. This is not to deny that such philosophical analyses of perception are "logically prior" accounts; rather, it is noted that science has proceeded quite well without waiting in the anteroom of philosophy for a definitive pronouncement on the issue, so the purely philosophical debate is duly noted and ignored. The psychologist takes up a certain burden by thus assuming the reality of reflexes, for it is then demanded of him that he be able to say of any aspect of any event that it does or does not belong to this or that part of a reflex, and that he be able to individuate reflexes from one another. These two tasks are complicated by certain experimental practices.

The specification problem appeared, for the first time, perhaps, in operant psychology, in a study by Reynolds (1961). Briefly, he reinforced responding in the presence of a white triangle on a red background,
but not in the presence of a white circle on a green background. The subjects were then exposed separately to the white triangle, the white circle, a red ground, and a green ground (all without reinforcement). Two subjects were reported that emitted between 45 and 55 responses per minute in the presence of the triangle and the red ground, but virtually none to the circle on the green (in the conditioning phase). In the extinction trials, one subject responded almost exclusively to the triangle, and the other to the red ground. Thus, while naive observations might have led one to believe that a discriminative stimulus for both subjects was to be described as "a white triangle on a red ground," it is clear that one subject's behavior was controlled by only the redness and the other by only the triangularity. Reynolds' method is a classical example of Mill's method of concommitent variation, and points to the necessity of applying this method whenever the precise specification of stimuli is at issue. The limitations of this method are well known; in Reynolds' experiment, the question may be asked whether the one subject actually responded to triangularity simpliciter, or to white triangularity, or angularity, etc. The multiplicity of possibilities at this simple level should give one pause as he contemplates such stimuli as the class of French Impressionist paintings, or even the class of chairs. The idea of actually specifying the physical properties of such classes boggles the mind.

The problem is solved for present purposes by recalling Boring's distinction between experimental and empirical psychology (Boring, 1950). For example, Brentano was explicitly empirical, but not experimental,
while his contemporary, Wundt, was both. The experimental problem is a formidable one indeed, but not impossible in principle. Interpretation, however, and thus the problem of concern here, is only an empirical activity. It is known that very wide classes can function as discriminative stimuli, and one simply assumes that there is some set of physical properties which define the class of things called chairs. The empirical approach does have limits, and considerable care must be taken in its use. It may be the case, for example, that some things are called chairs because they have one set of properties, and others because they have another set. Or, appellations by some may be due to some properties, while others may respond to other stimuli. Some of these problems may involve certain subtleties such as the notion of "open texture" developed by Waisman (1945). For the present, it will do to recall Skinner's characterization of interpretation in terms of "plausible guesses." Plausibility is a slippery term, but it remains the sole criterion for acceptable stimulus specification in interpretation.

A problem of response specification is raised quite elegantly by Catania (1973). His graphic representation of the problem deserves reproduction here. Fig. 1 shows the gradual selection of topography along one response dimension by the contingencies represented by the dashed line. Setting the parameter as time since the last reinforcement, the dashed line defines a fixed-interval schedule with a limited hold of \( y - x \). The original response distribution (A) is rather typical of a continuous reinforcement schedule, especially in early training. Distribution B shows the selective effects of reinforcement of only
Figure 1: The gradual selection of a response parameter from an entering repertoire (A), through an intermediate (B), to the final repertoire indicated by response distribution (C). The selecting schedule, a fixed-interval with a limited hold of \( y - x \), is indicated by the dashed line (after Catania, 1973).
those responses between $x$ and $y$, and $C$ shows a strong discrimination, probably unobtainable in pigeons without some sort of "external clock" (Ferster and Skinner, 1957). The difficulty that Catania points out can be seen by enlarging a distribution like $C$, as in Figure 2, which shows that there is a three-way ambiguity in the use that psychologists make of the term 'operant response class'. Researchers typically measure only the events generated by the contingencies ($B$), while they frequently speak of the class in terms of the contingency itself ($C$). Both practices ignore the events that are generated by, but do not meet, the contingencies ($A$). Catania's suggestion that the response class be defined as the class of events generated by the contingency is a promising one (Catania, 1973, p. 114). This approach incorporates the basic facts—the behavior and its "cause"—and still allows one to distinguish those events that meet the contingency and those which do not. (It can be noted in passing that contingencies are not always so easily specified as most researchers think. For example, if $y$ in Figure 2 were moved ten "units" farther out on the abscissa, it may or may not make any difference in the response distribution generated; certainly, it could be moved out far enough so that it made no difference. Catania's suggestion, then, makes sense only if the contingency used in the definition could not be changed without making some change in the behavior; that is, the functional contingency must be specifiable.)

Catania's definition leads to several consequences. It follows that topographically identical events belong to different classes if they are generated by different contingencies, as well as that topographically
Figure 2: A response distribution compared to a fixed-interval with a limited hold contingency. Area (A) consists of responses which do not meet the contingency; (B) of those that do; and, (C) of possible responses which would meet the contingency but do not occur (after Catania, 1973).
distinct events fall in the same class if they are generated by the same contingency. Of course, contingencies generally specify topographical limits. Except in such cases as "superstitious" behavior (Skinner, 1948) and such classes as "creativity" (Pryor, 1969), then, operants are defined in terms of topography and controlling variables.

With this definition in mind, a problem of individuation of operants generated by the phenomena of "stimulus generalization" and "shaping" is easily solved. The generalization phenomenon is illustrated by a pioneering study by Guttman and Kalish (1956). They trained birds to peck a circular key transilluminated by one color (i.e., one wavelength, or very few indeed), and then measured response rates in extinction to other colors presented randomly for short durations; included in the random test was the original training stimulus. The results of this kind of procedure are represented in Figure 3, where \( x \) indicates the color trained to, and \( y \) and \( z \) mark arbitrary points at higher and lower wavelengths. One might be tempted to say that this procedure generates an indefinite (or infinite!) number of operants, one for each of the discriminable real values of wavelength in the visible spectrum, or each possible color. However, the behavior is generated by exactly one contingency: training to \( x \), and extinction under the various colors. The control of pecking by \( y \) and \( z \), actually, is not established by the testing sequence, any more than a magnetized nail's attraction to ferrous metals is established by placing it near such a metal and seeing that it is attracted. The issue raised by generalization, then, is solved by counting contingencies, which can
Figure 3: A generalization gradient. The wavelength reinforced in training is $x$; $y$ and $z$ mark arbitrary wavelengths other than that trained to (after Guttman and Kalish, 1956).
be done using Catania's graphing method, expanded trivially to include enough abscissas to represent all of the properties of the response class which are specified by the contingency. A single n-dimensional dashed line then represents a single contingency.

Shaping, on the other hand, is a technique which consists of gradual alterations of contingencies, thus gradually generating responses with different topographies via different contingencies. The process can be illustrated using Catania's method, as in Figure 4. Let the response parameter be force of bar press; contingencies A–E represent successive increases in the force required of a response to trip the reinforcement delivery mechanism, and response distributions 1–5 represent distributions generated at six different moments. It is clear that the actual facts would demand an indefinite number of response distributions, for, although the contingencies A–E may be changed digitally, the data indicate a more nearly analogue shift in response distributions under the proper and typical conditions. Two problems seem to be generated by this phenomenon. Suppose that contingency A is in effect, and that distribution 1 represents a stable performance that A has generated. Now, upon changing to contingency B, an indefinite number of distributions will be generated before 2 is obtained, which can be supposed to represent a stable performance under B. What of all the distributions generated between 1 and 2? How many operant classes were generated by B? There must have been a distribution 1½ generated sometime between the appearance of 1 and 2, and one cannot but ask whether it is identical with, or belongs to the same response class as 1, 2, or neither. A related difficulty
Figure 4: Successive response distributions (1–5) generated, respectively, by successive contingencies (A–E) (after Catania, 1973).
is to decide which of the responses in $l_{1/2}$ are due to A and which to B, given that $l$ is not a stable performance under A.

There seems to be no escape from the conclusion that these problems force us to conclude that many members of an operant class belong simultaneously to two response classes; or, alternatively, one might conclude that there is but one class, belonging to more than one operant class, corresponding to more than one contingency. Finally, in an effort to retain our intuitions or habits regarding response and operant classes, we might opt to redefine contingencies in terms which allow us to speak of the shift in shaping as belonging to only one contingency, one defined in terms of rates of change over any dimension(s).

Perhaps it is this last alternative that must be embraced. Science has often analyzed an essentially continuous process into static parts, like a movie, where each frame takes a slice of time, but each successive pair of frames leave something of the process out. Even Newton's solution in terms of the calculus of infinitesimals suffers from the same inadequacy, for delta-$t$ can be taken as small as you like, but delta-$t$ remains a change in time, not an instant; a true instantaneous velocity, for example, requires one to calculate for delta-$t$ equal to zero, not merely approaching zero. Skinner recognized this sort of problem early in his career (Skinner, 1931; 1935a), and saw it as an inevitable result of analysis. Indeed, it is not surprising that, when forced into parts, the continuum of the phenomena force us to place certain "events," themselves abstractions, into two frames of our conceptual movie.
There are reasons to choose the two-class solution, and others for choosing the redefinition of schedules in terms of changes. Perhaps a way between the horns can be found by regarding $1^{\frac{3}{4}}$ (and $1^{\frac{1}{4}}$, etc.) in Figure 4 as stages in a process: the process of the generation of the operant class defined in terms of B and 2, assuming here that B,2 define a steady state. It may prove useful, empirically or experimentally, to distinguish operant B$^{1\frac{3}{4}}$ from B,2, in which case a new approach to the problem may be needed, perhaps involving new definitions. At this point, however, the changes in behavior that result from the imposition of new contingencies may be looked upon as a process of formation. An odd result of this point of view is that only stable performances can be properly individuated into operants, unless responses occurring during dynamic changes are viewed as belonging to more than one class, defined by the previous contingencies and the new ones. The former view being unacceptable, we seem pushed again to embrace response-instances belonging to more than one class. Mill's method of residues can be utilized to ascertain which of the instances in a distribution are due to a previous contingency, and which to the new, thus providing a measure of the degree to which a distribution, or response class, is due to the first or second contingency. And this, in turn, provides a measure of the progress of generating a new operant class.

Summary

We have examined the basic terms of the operant analysis and found for each of the major terms a rather clear grounding in the
physical. Certain problems were raised regarding the way to describe the phenomena of generalization and shaping. Some of these problems were found to be most troublesome, but it is encouraging to see that they appear in similar form in other sciences. No final solution was offered, though suggestions were made.

The basics, then, have been covered. A form of the intentionalist challenge, in terms of purpose and content, was met, and the implicit ontology of eliminative materialism has been explained and defended.
CHAPTER VII

Summary and Conclusions

The basic argument, throughout, has been that traditionally mental language, insofar as it seems to demand acceptance, describes nothing but operant phenomena, or aspects thereof. The argument of Chapter VI was to show that the language of operant behavior reduces to talk about physical events and relations among them. Together these conclusions were used, in Chapter V, to show that the eliminative materialistic program of Skinner's radical behaviorism is a plausible one.

The net effect, then, was to argue that mental or intentional language can be, in principle, eliminated in favor of language which refers only to the commonsensically physical realm, and relations in that realm. The specific intentional idioms considered were those of purpose, content, awareness, and belief; however, these were chosen because they are so central to the opposition theses. Insofar as the argument succeeds in reversing the system of failures mentioned by Davidson, it succeeds also in establishing the radical behaviorist's program as a plausible one.

One obvious implication for philosophy is the serious reconsideration of the behavioral approach. In particular, behaviorism as a reinforcement theory should be reconsidered, for it is the role of reinforcement, as Skinner noted (1935b), that is crucial to an interpretation of purpose. Traditional philosophical discussions of behaviorism, virtually without exception, deal only with inputs (stimuli)
and outputs (responses), thus missing Skinner's fundamental contribution of a distinction among three types of stimuli: the respondent, the operant discriminative stimulus and the operant reinforcing stimulus. The utilization of the reinforcement concept should be made easier for the philosopher by the symmetry shown between it and the ordinary language concept of purpose. There can be no doubt that reinforcement is operative in much of human behavior, even if the ubiquity claimed by Skinner is doubted. And, I think, there can be no doubt that the ordinary concept of purpose displays an extraordinary symmetry with reinforcement. It seems too much for coincidence that this symmetry should exist without there being the obvious causal connection: the ordinary concept developed as a description of the facts scientifically described by reinforcement.

Another implication for philosophy concerns Brentano's thesis. We have found the traditional area of intentionality to correspond to the area of operant phenomena. The reinforcement relation, the essence of the operant, is unparalleled in nature: feedback systems abound, but they are no more similar to the operant reinforcement relation than is the respondent relationship between response and stimulus. Thus, there is a formulation of Brentano's thesis that remains unchallenged by anything said here, although the contrast with the physical must be abandoned: there is an area of phenomena called the operant, which is fundamentally different from other natural phenomena. And, given the present analysis, this new version of Brentano's thesis separates the same phenomena from all others, albeit not
from the "physical."

An implication for psychology is to be found in the successful reduction of the mental language to operant terminology. Insofar as the traditional language can be said to refer to nothing but operant phenomena, it can be said to refer after all. Thus, the traditional language should not be rejected as "meaningless," as many behaviorists would have it; rather, a proper view would be that, for any given reliable traditional idiom, it very well may describe an important aspect of behavior, an aspect important enough to be discriminated over the years. And to that extent, it may well be an aspect worth investigating. The ordinary mental language, then, should be dealt with in respect that is partial or tentative if not total.

A field of applied psychology that may especially benefit from this new orientation is that of behavior modification as applied to what may be called, for want of a better term, garden variety neurotics. When an organism's repertoire is not noticeable deficient to the ordinary intelligent observer in any specifiable way, but that organism persists in "being unhappy," our technology is often inadequate. Not that we do not know how to change the behavior of such persons. Rather, the problem is that we do not have any special skills or knowledge that enable us to help define the direction of change that would be more/most desirable. Compare with this the wealth of definitions, offered in traditional language, of "normal" or "healthy" repertoires. Perhaps an energetic attempt to interpret some of these definitions of how a person should behave if he/she is to be "happy" would supply working
hypotheses at least for clinicians to investigate.

A more theoretical and more basic implication is to be found in the interpretations offered herein for intentional idioms. Without exception, they depend upon and refer to that portion of an organism's history which is responsible for a part of his present repertoire. It is hard to see how to avoid this sort of reference, and, yet, it is hard to find any research in the literature which substantiates such reference. That is, we have no knowledge of how to tell, given an entire history, which part is causally relevant to which part of the current repertoire. Nor is this of interest only because it is needed for present purposes. Although operant studies are notably more longitudinal than most others in psychology, they are short compared to the organism's lifetime. When we know that a given schedule (history) yields a given pattern of behavior, we would also like to know what a slightly different schedule would yield. And, ultimately, we would like a theory which predicts what small alterations in a schedule would do to the terminal repertoire. Taking this line of reasoning one step further, we would like to be able to predict the effect of a small change in one of many concurrent schedules on any or all of the various responses maintained by those concurrently effective schedules. And this predictive ability would amount to nothing less than an experimental ability to say, from a given history, which events contributed to which parts of the present repertoire. Thus, the present discussion of the intentional does not so much raise this issue as to bring it into sharper focus.
REFERENCES


Hull, R. T. On taking causal criteria to be ontologically significant. *Behaviorism*, 1973, 2, 65-76.


