

Do Stimuli Elicit Behavior?—A Study in the Logical Foundations of Behavioristics

Abstract

It has become customary in modern behavioristics to speak of stimuli as though they elicit responses from organisms. But logically this is absurd, for analysis of the grammatical roles of stimulus and response concepts shows that stimuli and response differ in logical type from causes and effects. The “S elicits R” formula thus stands revealed as elliptical for a more complicated form of assertion. The trouble with this ellipsis, however, is that by suppressing vital components of formal structure in behavioral principles, it has led to gratuitous assumptions about the environmental antecedents of behavior and seriously undermined the ability of behavior theory to assimilate the “higher mental processes.”

A hard-shelled empiricist at the time of the Copernican revolution in astronomy (and there surely were such creatures, even in those days of methodological innocence) must have found it difficult to appreciate what the shouting was all about, or at least could have seen little excuse for it. For apart from theological overtones and computational ease, there is really no practical or even theoretical difference between heliocentric and geocentric interpretations of celestial motion. Mathematically, one can always transform one frame of reference into the other, and a suitably elaborate Ptolemaic description of planetary movements is equivalent to the heliocentric description. Yet the Copernican shift in viewpoint made possible the discovery of Kepler’s laws, without which Newton’s labors would never have reached fruition.

*

Recital of an episode from the history of astronomy may seem a peculiar way to introduce a methodological issue in contemporary behavioristics, but I wished to illustrate how the development of a science may depend crucially upon a conceptual shift which in itself contributes no new factual material. For at first impression, the point I am about to make, concerning as it does certain niceties in the grammar of behaviorese, may seem to be merely a hair-splitting illustration for Adams’ (1954) gibe that “methodology . . . is a sterile and ineffectual business.” I wish to call attention to certain logical anomalies in our way of talking about the relation of stimuli to behavior; and because the argument is purely a formal one, the empirically oriented reader’s first reaction is likely to be, “So what?” Yet I feel that

certain clarifications in our conceptualization of the external causes of behavior are essential if we are to penetrate much further into the mysteries of the behaving organism.

Briefly, the point to be made is that stimuli are *not* causes—not even part-causes—of behavior, for the simple reason that stimuli are not logically the kind of things which *can* cause anything. That is, the expressions by which we describe stimuli are of a different logical type from those by which we describe causes. The reason why this is of methodological importance is that when we say such things as “This response was elicited by stimulus *S*,” or “*S* is a discriminatory stimulus which set the occasion for this behavior,” we *are* using stimulus-terms as though they referred to part-causes of the behavior. And while this is in itself a relatively harmless ellipsis which, as a rule, is in no way misleading in regard to what transpired during the occasion to which reference is made, it establishes certain unwarranted preconceptions and misguided habits of thought which seriously if not hopelessly impede understanding of the more complex forms of behavior. In particular, it has led to the behavioristic use of simple terms where expressions of greater logical complexity are necessary (a situation crudely illustrated by an attempt to represent forces in physics by single numbers rather than by vectors in 3-space), with resultant failure of behavior theory to give satisfactory account of such phenomena as “relational” responding and cognitive processes.

The Grammar of Elicitation-Talk

What is a stimulus? This is a question which has been asked many times by analytically minded psychologists. What is ordinarily at stake in such queries are certain empirical facts about the organism and its environment. Here our problem is of a different sort: What *logical kind of* an entity is a stimulus? Or less esoterically, into what grammatical category do the expressions by which we refer to stimuli fall? This is simply a question about the technical language of behavioristics and is answered by examining representative statements by reputable behaviorists in which such expressions are used.

It would seem first of all that the concept of “stimulus” has been applied at two different stages in the sequence of behavioral activation—to aspects of the *environment* (“A part, or modification of such a part, of the environment is traditionally called a stimulus” (Skinner, 1938, p. 9)), and to aspects of the *activity of sense receptors*. Now the latter, while perhaps closer to the historical origins of the concept, are physiological variables whose values are unobserved, or even inaccessible to observation, in the vast majority of behavioral studies. Modern behavioristics, on the other hand, with its increasing attention to empirical phenomena, has come to regard “stimuli” primarily as part of the datum-level antecedents of behavior (although theoreticians frequently subsume certain hypothetical processes under

this term as well). Hence, we may confine attention to the first case, although the conclusions which will be drawn would not be changed if we also included the latter. However, this still leaves undetermined what *sort* of environmental “aspects” stimuli are. Examination of the literature reveals that psychologists have habitually identified as stimuli entities of two logical types: (a) *objects* (e.g., individual cage-floors, particular water-cups, specific food-pellets), and (b) *attributes*¹ (e.g., red, hard, triangular), or thing-kinds (e.g., dog, food, stick), which may here be regarded as attribute clusters or disjunctions of behaviorally equivalent attribute-clusters. For example, Woodworth (1921, p. 486f.) identifies a great number of objects as stimuli for play behavior: “Thus we get the following classes of play-stimuli . . . things [such as] bicycle, skate, sled, . . . kites, rockets, boats, balls [etc.]” It is plain that since a class of objects is identified as a class of stimuli, the individual stimulus is here conceived as the individual object. Hilgard and Marquis list, as possible conditioned stimuli, “lights of various colors, papers, geometrical forms, rotating objects, . . . pure tones, horns, buzzers, . . . metronomes, electric fans, . . . flashes of light, clicks” (1940, p. 35f.), while Skinner identifies the stimuli in a hypothetical experiment as “food,” “tray,” “sound,” “visual lever,” “tactile lever” (1938, p. 52f.). The latter examples list a mixture of objects and attributes. Thus colored lights, rotating objects, buzzers, are specific objects, whereas geometrical forms, pure tones, lightflashes, sound, tactical and visual lever² are best interpreted as attributes of objects. Stimulus-terms such as ‘food’ and ‘tray’ are ambiguous in that they do not make clear whether the stimulus is the specific tray-like or foodlike object, or the set of attributes in virtue of which the object belongs to the tray or food thing-kind. Apparently, we are accustomed to regarding both the object itself and its attributes as (possible) stimuli, and to be sure, there is little to choose between saying that an object is a stimulus in virtue of a certain attribute, and that the attribute is itself the stimulus. However, broader contexts in which the notion of “stimulus” is invoked indicate clearly that the basic use of this term is to denote attributes. Suppose, for example, that n presentations of a card bearing a red circle, followed by electric shock, have established a conditioned GSR in some organism o . Even if o has been shown a different card with

¹It is not really legitimate to lump all “attributes” into a single logical type, since there are important logical differences between, say, attributes of single objects (e.g., red), relational attributes (e.g., being-larger-than), attributes of attributes (e.g., intense), etc. For present purposes, fortunately, these differences may be ignored. Neither are we here concerned with the difference between the adjectival (e.g., “circular”) and singular (e.g., “circularity”) forms of predicates, important as this distinction may be for problems in ontology. In fact, even the difference between objects and attributes is basically irrelevant for the problem at hand. What is crucial here is the logical distinction between *terms* and their referents on the one hand, and *sentences* (or structurally equivalent expressions) and what they describe on the other.

²We may construe such things as noises, visual and tactile lever, etc., as attributes of the physical objects involved in the same way that we regard, say, red as a property of ripe apples rather than as an excitation of our visual receptors.

the red circular marking on each trial (as might be done to eliminate extraneous cues), it would still be said that *o* had received *n* presentations of the conditioned stimulus, never that *o* had received one presentation each of *n* different stimuli. To take another example, certain regularities found in reflexive behavior are described by saying that the latency of the response is a function of the intensity of the stimulus. But the concept of “intensity” (or degrees thereof) applies only to attributes, not to objects—a fire truck, e.g., is not itself intense, although it may very well be an intense red in color. Hence by “stimuli,” we seem primarily to mean certain attributes of things, although we sometimes also apply the term to individual objects.

In like manner, we need to determine what logical kind of an entity may correctly be said to elicit behavior, or to increase the likelihood of the occurrence of a certain response. We no longer need specifically to consider the language of behavioristics for this purpose, since what we are now asking is simply what kind of an entity can *cause* another—where for want of a better term, we are using ‘cause’ not in its exciting but hotly disputed and never satisfactorily analyzed philosophical sense, but merely in the relatively innocuous scientific sense of something from which predictions can be made (or perhaps, a poorly defined subclass of such). That is, we are asking what kind of an expression can occur in place of ‘*A*’ in grammatically correct sentences of form ‘*B* can be expected with greater confidence when *A* is known than when *A* is not known,’ ‘*B* is predictable on the basis of *A*,’ or, more idiomatically, ‘Since *A*, it is likely that *B*,’ or ‘*B* because *A*.’

It does not require much reflection to see that causes and effects—i.e., that from which and to which predictions are made—can only be states-of-affairs, that is, *events* or *facts*, by which terms we mean what it is that makes a sentence true. Thus it makes perfectly good sense to say, ‘Since the sky is now dark with clouds, it is likely that it will rain tonight,’ or ‘Mary slapped John because John pulled Mary’s pigtail.’ In the first instance, we imply that because the sentence, ‘The sky is now dark with clouds,’ is true—i.e., because it is a *fact that* the sky is now dark with clouds—it is reasonable to suppose that the sentence, ‘It will rain tonight,’ is also true—i.e., to suppose it to be a fact that it will rain tonight. Similarly, in the second example, we imply that once we know that John pulled Mary’s pigtail, we have grounds for belief, even before being informed of Mary’s response, that the sentence, ‘Mary slapped John,’ is likely to be true. In short, scientific prediction is a judgment, based on what we believe to be the case, about what is therefore also likely to be the case. Actually, once we recognize that for science, causes are *grounds for prediction* (or at least a subclass thereof), and that prediction is a form of *inference*, it becomes clear from grammatical considerations alone that a cause must be something which is signified by a *sentence*, for inference, whether formalized according to strict rules of logic or not, is the derivation of one proposition from another.

(It is true that expressions which suggest that causal efficacy obtains between entities other than facts are not uncommon in idiomatic English. For example, we say, “A worm caused this apple to spoil,” or “Heart-failure is a cause of death,” or “John’s broken leg is due to his having been in an automobile accident.” But a little analysis shows these to be ellipses or variants of straightforward propositions expressing a relation between facts. Thus the first example is shorthand for something like “This apple spoiled because a worm burrowed into it,” the second is elliptical for a lawlike statement such as “If any person x has heart-failure at any time t , then x is dead at time $t + 10$ min.,” while the third is an event-talk paraphrase of “The fact that John broke his leg is attributable to the fact that he was in an automobile accident.”)

By now it should be obvious why I claimed earlier that stimuli cannot be causes of behavior. For stimuli are *attributes* (or occasionally objects), whereas a particular response-occurrence is a fact which can be caused only by another *fact*. The only way in which a stimulus can share causal responsibility for a behavioral event is by being a constituent of some fact which is a causal antecedent of that behavioral event. Thus the claim, “This response was elicited by stimulus S ,” or “ S is a discriminatory stimulus which set the occasion for this behavior,” if taken literally, is nonsense—the proper formulation in something like, “This response-occurrence was elicited by *an occurrence* of stimulus S ,” or “This fact about S set the occasion for this behavior.” Actually, the latter modes of formulation, or something very like them, are by no means infrequent in the literature. Thus Skinner, who has no qualms about saying that a stimulus elicits a response, resorts to event-terminology whenever he wishes to be more explicit; e.g., “The term [‘elicitation’] describes the fact that *presentation of* a stimulus is followed by a response.” (Skinner, 1938, p. 234; italics added). In general, the literature makes abundantly clear that the expressions, ‘stimulus S ’ (which, taken literally, denotes attributes) and ‘occurrence of stimulus S ’ (which denotes events), are used more or less interchangeably. Hence, few behaviorists should have reason to dispute that it is not the stimulus in and of itself, but at best the fact that it occurs in the organism’s immediate environment at a certain moment, which is (in part) responsible for the organism’s behavior shortly thereafter.

Elicitors and “Structure”

Now, it is one thing to accept this apparently innocuous grammatical point, and entirely another to appreciate its significance. For the behaviorist’s retort is likely to be: “When I say ‘This behavior was elicited by stimulus S ,’ I *mean* that the behavior was elicited by an occurrence of S . All I am doing is to adopt a convenient abbreviation which I can always expand whenever necessary.” Now this is quite

true so far as it goes. Whenever it is the case that a given bit of behavior is elicited³ or predisposed by the occurrence of the stimulus, it is harmless to speak as though the stimulus were itself the elicitor. What I am concerned about is something quite different from, and much more serious than, the grammatical refinements of behaviorese. Traditional S-R jargon, in which a stimulus is spoken of as though it were in itself the elicitor of the response, plus failure to recognize *explicitly* that the imminent external cause of a behavioral event can only be another event of which the stimulus is but an ingredient, *have induced behaviorists unwittingly to rule out as possible environmental elicitors of behavior all but a restricted class of events, limitation to which is justified by neither general methodological considerations nor empirical observation.* For it is but a short leap—almost irresistible unless one knowingly refrains from making it—from regarding behavioral phenomena which can be described conveniently by the elliptical form, ‘This response was elicited by stimulus *S*’ as the *paradigm* case of behavior-elicitation, to assuming that *all* states of the environment to which an immediately following behavioral event is (in part) attributable should be so describable. We then find ourselves saddled with the wholly gratuitous assumption that only facts for which we have available descriptions roughly of the form, ‘Stimulus *S* is present in the immediate environment of organism *o* at time *t*’ may be regarded as elicitors of behavior.

Now, it is palpably false that all environmental events which we have good reason to believe have evoked behavior are describable, in any readily discoverable or formally simple way, by conjunctions of sentences of the form just indicated. Thus we have strong evidence that organisms are conditionable to *relational* facts, a finding which has aroused great consternation among orthodox S-R theorists, and has led some to deny that the response is *really* influenced by the relational element. The classical example, here, is that of Köhler’s chickens, which presumably learned to go to the darker of two stimulus-cards. Another interesting instance is Robinson’s “oddity” problem, in which monkeys learn to select from three objects, two of which are alike, the one which differs from the others. Faced with such intransigent brute data, some behaviorists have tried to find in the eliciting situation special “stimulus-configurations,” sets of competing stimuli, or the like, to which could be correlated response-tendencies in a way that would yield prediction of the observed behavior. Now, I do not doubt that given a sufficiently elaborate concept of “stimulus,” a sufficiently complex theory of generalization gradients, etc., most if not all environmental events to which a response elicitation is attributable can, with sufficient ingenuity and patience, be given what at least *seems* to be a mere-presence-of-stimuli kind of description, any more than I doubt that planetary motions can be described by a sufficiently complex set of Ptolemaic cycles and epicycles. The primary consideration in such matters, however,

³Henceforth, I shall use the term ‘elicitation’ in that broad sense (cf. Skinner, 1938, p. 241) which also subsumes the discriminatory, or “cue,” action of stimuli.

is not whether the phenomenon *can* be formulated according to a certain conceptual frame, but rather, which formulation is the most illuminating with respect to underlying principles.

There is no need here to review the struggles of orthodox behavior theory in its efforts to cope with elicitations which appear to depend on a relational element. It suffices to note that responding of this sort has posed much difficulty for the behavior theorist. We have ample historical evidence that a conceptual frame which can be imposed only with a great deal of grunting and straining on certain commonplace phenomena which logically fall within its scope invariably proves to be a scientific dead-end. The trouble in the present instance, however, lies not with an elicitation theory of behavior (which, broadly conceived, is not a theory at all, but merely a generalization from known data), but only with the distortion unnecessarily introduced by speaking of the stimulus-components of the eliciting fact as though they were themselves the elicitors. How this distortion creeps in may be appreciated by considering that as alternatives to the elicitation-by-“absolute”-stimuli interpretation of responses to relational facts (i.e., the theory that the response of an organism which apparently learns to go to the darker of two cards is in reality the algebraic resultant of individual response-tendencies to the various non-relational attributes present), we find proposed (*a*) the hypothesis that the organism responds to the relationship itself, and (*b*) the view that the organism responds to the “total stimulus-configuration,” or the “situation as a whole” (see Spence, 1942). Interpretation (*a*) results immediately from retaining the habit of speaking about the dominant element in the eliciting event as though it were itself the elicitor. But to say that the organism responds *to the relationship*—e.g., to larger, darker, etc.—simply fails to be at all a satisfactory description of what is going on. The bird which chooses the darker card is not responding to being-darker-than—i.e., to the fact that the relationship, being-darker-than, is present in the environment—but (presumably) to the fact that card *x* is darker than card *y*; the mere presence of being-darker-than does not suffice to determine the response (see below).

As for (*b*), the situation-as-a-whole interpretation, while it may thrill one’s poetic soul to contemplate the Unity of Nature, the job of the scientist is not merely to nod his head sagely as data pass by, but also to make some reasonably accurate predictions about those data—which he can do only by carving Nature into pieces, shriek and bleed though she may. But denial that the behavior-eliciting powers of a situation can be exhaustively described by listing the various attributes, whether “absolute” or relational, present in the situation, does not entail the lethal belief that the totality of those aspects of the situation relevant to the elicitation cannot be succinctly encapsulated by a formally rigorous technical language—all we need is a satisfactory way to record the *structure* of the eliciting event. Now the term “structure” has an unsavory aroma for the tough-minded behaviorist,

for it has too frequently been used to signalize the difference between the fragments of a situation and the ineffable wholeness of it all. But actually, we have a perfectly sound symbolic mechanism for recording structures—namely, *sentences*. The “structure” of an event is merely the difference between what is described by a list of words and by a sentence comprised of those words. That is, a set of objects and attributes together with their structure is simply the *fact* of which these entities are ingredients. (Thus John, Mary, and the relationship of loving might be constituents both of a fact described by ‘John loves Mary’, and also of another fact described by ‘Mary loves John’. The only difference between the two facts is one of structure, yet their causal ramifications may be highly dissimilar.) To claim that an organism responds not to the discrete elements of the situation but to the situation-as-a-whole, is either an obscurantistic denial that a science of behavior is possible at all, or is merely a confused way of recognizing that facts of which objects and attributes are ingredients, not the objects and attributes in themselves, are what elicit behavior. The connotations that “holism” has acquired (not wholly unjustly) for rigor-oriented behaviorists illustrate nicely how misleading it is to speak of stimuli as though they themselves evoke responses. For the S-R theorist would normally regard a conception of behavior which professes to take into account the “structure” of the environment as a rather vague alternative to his own theory. But since only facts can be causes of behavior, there can be no structureless elicitors; there is only the counter-intuitive possibility that only a special kind of structure—namely, that which characterizes facts of the mere-presence-of-stimuli kind—can play a role in the instigation of behavior.

It is not practical here to analyze in detail the importance of being clear that only facts (or events) can elicit behavior, or to dissect actual instances of major psychological issues on which our current manner of speech has managed to snarl effective thinking. What is at the root of these vicissitudes of contemporary behavior theory is a methodological inadequacy: By trying to make simple *terms* do the work of expressions which need the formal structure of a *sentence*, behaviorists have adopted a conceptual framework which, entirely apart from the meanings of the concepts employed, lacks sufficient *formal complexity* to cope successfully with more intricate psychological phenomena such as perception, semantical processes, belief, etc. But correlatively, these are precisely the phenomena to which one must turn in order to demonstrate this formal deficiency most convincingly, and they are not matters which can illuminatingly be dealt with in a few paragraphs. Still, abstract generalities are not always by themselves very convincing, and it yet remains to be shown that behind the shadow of the present remarks lies the raw substance from which real scientific advances are hewn. Hence to support my argument that expressions by which elicitors are formally represented in the language of behavior theory need to have the syntactical complexity of a sentence, I shall try to illustrate how the structure of a stimulus-event can make a difference,

additional to the import of its constituent stimuli, for the elicited behavior.

It would be highly mistaken to assume that insistence upon explicit recognition that only a fact can evoke a response finds practical justification, if any, only in recondite theoretical analyses of complex behavior. Quite the contrary, recognition—if not explicit, then of necessity implicit⁴—that an elicitor is something more than just a set of stimuli emerges on the grass-roots level of empirical observation, quite apart from any theoretical or methodological speculations. To appreciate this, we need only to examine the logical details of the data-language descriptions of behavioral regularities as they are actually observed.

As Skinner (1938, p. 9) has pointed out so forcibly, the correlations of certain features of the environment with certain aspects of an organism's activity are simply part of the brute data of behavioristics. Thus we might summarize a laboratory protocol, after description of the experimental arrangement, etc., by "On April 6, 1960, sounding of the bell was, in 96% of the cases, followed within five seconds by increased salivation from dog #8." On the basis of such an observed regularity, we make the inference (and it *is* an inference, rather than just a summary of the data, and not merely a statistical generalization either), expressed in S-R jargon, that dog #8 has a salivation-to-sound-of-bell habit; and we would explain what we meant by a "salivation-to-sound-of-bell habit" by saying that sounding a bell in the presence of an organism with such a habit will cause it to salivate, or, more explicitly, that if an organism o has a salivation-to-sound-of-bell habit at time t , and a bell is sounded in o 's presence at time t , o will, with high likelihood, show an increase in salivation within (say) five seconds after t . The methodology of habit-concepts is here irrelevant, however, for our present concern is only with description of the empirical regularities on which they are grounded. If 'C(o,t)' abbreviates description of an observable condition under which we would be willing to infer that an organism o has a salivation-to-sound-of-bell habit at time t (such a condition being, e.g., that o salivated on the last fifty trials prior to t), we have an empirical generalization wholly in observational, or data language, terms: "If C(o,t), then if a bell is sounded at time t in o 's presence, o will (very likely) salivate within five seconds after t ." The point to be made here is that behavioral psychology is grounded on a large number of direct statistical generalizations from raw data to the effect that if certain observable conditions obtain for an organism o at time t (e.g., a certain conditioning history, certain facts about past performance, etc.), then, if the immediate environment of o at t conforms to certain specified and observable characteristics, o will usually behave in a specified way shortly thereafter. The relevance of this for present purposes is that in such empirical generalizations, description of the elicited behavioral event frequently makes use not only of the individual ingredients of the eliciting environmental event, but

⁴Implicit efforts to capture this additional formal structure are illustrated by concepts such as "compound stimulus," and especially, "stimulus configuration."

also of the *way* in which the elicitor is put together. I illustrate this with two commonplace examples.

Suppose we give an organism repeated trials on which it is strongly reinforced for going to a red stimulus-card. This is a typical discrimination experiment, the details of which the reader can easily fill in for himself, and many years of animal research have taught us that a mature, healthy animal of reasonable intelligence, if exposed to a sufficiently long series of such trials under sufficiently diverse circumstances (the latter to rule out extraneous cues), will acquire a habit of a kind frequently described as “responding positively to the stimulus,” where the stimulus is here the compound attribute of being both red and card-shaped. If a person unacquainted with S-R jargon asked what we mean by “responding positively” to being-a-red-card, we would tell him in ordinary language that it means having a tendency to approach red cards. If ‘ $C(o,t)$ ’ now abbreviates the complex statement which asserts that at time t , organism o has had the conditioning history just indicated, and ‘ $H(o,t)$ ’ asserts that o is healthy, etc., at time t , we may then describe our empirical generalization by the statement:

- (1) If $C(o,t)$ and $H(o,t)$, then, if x is a red card in the immediate environment of o at t , o (likely) moves toward x at t .⁵

For our second illustration, let “ $D(o,t)$ ” describe a training procedure similar to that described by “ $C(o,t)$ ” except that instead of reinforcing positive responses to all red cards, if there is more than one red card present on a given trial, only going to the largest is reinforced. The results of the latter conditioning operation are not so easily described in S-R jargon, since this is an instance of relational responding, yet the *data language* account is simple enough—we simply say that the animal learns to go to the largest red card present. The fully expressed generalization in this second case is:

- (2) If $D(o,t)$ and $H(o,t)$, then, if x is a red card in the vicinity of o at time t and there is no other red card in the vicinity of o at t which is larger than x , o (likely) moves toward x at t .

It will be observed that in both (1) and (2), the elicited behavioral event is described by a sentence derived from the schema

‘ o moves toward x at t ’.

The contrast between this formula and one like

⁵In a more sophisticated formulation, we would say that o has a *tendency* to move toward x at t , where the concept of “tendency” assimilates those cases where elicitors of competing responses are co-present.

‘*o* raises his right *paw* at *t*’

is important. A “response” is something done by a specific organism at a specific time. Hence a sentence-schema about behavior, such as ‘*o* raises his right paw at *t*’ is indefinite with respect to (i.e., contains variables for) *only* the organism and the time is a full description of a response, since one need fill in only the name or description of a definite organism and time in order to get a (possibly false) statement about the organism’s behavior at that time. Conversely, a sentence-schema about behavior which contains additional variables, such as

‘*o* flexes his *y* at *t*’,

in which ‘*y*’ ranges over anatomical parts, does *not* fully specify a response, for we ascribe different responses to an organism *o* at time *t* according as we substitute ‘left foreleg’, ‘tail’, etc. for ‘*y*’. Similarly, ‘*o* moves toward *x* at *t*’ does not specify a fixed response; such a description is obtained only by substituting for *x* an expression (e.g., the name or description of an object or a position in space) such that when specific names are then substituted for *o* and *t*, a complete, meaningful, sentence is obtained. The significance of this is that generalizations (1) and (2) *do not specify a fixed response elicited by the stimulus-event*. In order to predict what behavior is elicited—e.g., going-to-object-*a* vs. going-to-object-*b*, etc.—we have to know the details of the eliciting event, where these differ from one occasion to the next as *x* assumes different values. Moreover, to predict the response, we not only need to know what elements are present in the elicitor, *we also need to know how they are combined therein*—i.e., how they are “structured.” Thus the information that object *a*, the attribute of being red, and the attribute of being card-shaped are all present to an organism *o* at time *t* is inadequate for prediction, by (1), that *o* likely moves toward *a* at *t*. We also need to know that being red and being card-shaped are both attributes of *a*, rather than of other objects present to *o* at *t*. The essential role of structure is even better illustrated by (2). For the sentences, ‘*a* and *b* are the only red cards present to *o* at *t*, and *a* is larger than *b*’, and ‘*a* and *b* are the only red cards present to *o* at *t*, and *b* is larger than *a*’, describe (possible) events whose ingredients are the same, and whose only difference is that the elements are combined differently—i.e., in the one, *a* is larger than *b*, while in the other, *b* is larger than *a*. Yet for the one event, the prediction is that *o* will go to *a*, whereas for the other, it is that *o* will go to *b*.

Now it might be protested that these examples support my argument about the behavioral importance of the eliciting event’s structure only because of the way I have chosen to word (1) and (2). For example, it might seem that (1) could be rewritten as:

(1’) If $C(o,t)$ and $H(o,t)$, then, if there is a red card in the immediate environment of *o* at *t*, *o* (likely) moves toward a nearby red card at *t*.

As described by (1'), the response elicited by the presence of a red card is always the same—namely, moving-toward-a-nearby-red-card. Similarly, in (2) we might try replacing 'o moves toward x at t' with 'o moves toward the largest red card present to o at t'. In fact, entering into the spirit of things, we might as well speak of a red card which is larger than any other near o at t as having the attribute, being-a-red-card-larger-than-any-other-nearby-red-card, giving us a reformulation in the classical S-R tradition in which a fixed response is elicited by a mere-presence-of-stimulus kind of event:

(2') If $D(o,t)$ and $H(o,t)$, then, if the attribute of being-a-red-card-larger-than-any-other-nearby-red-card is present to o at t, o (likely) moves toward the largest red card present to o at t.

There are various considerations which can be brought against acceptance of (1') and (2') as adequate reformulations of (1) and (2). (For example, if two stimulus-events which elicit the same response are co-present, we should expect the tendency to that response to be enhanced. But in the case of (1'), if both a and b are red cards on opposite sides of o at t, we have an approach-approach conflict whose net effect is to weaken the tendency to move-toward-a-red-card.) However, even if (1') and (2') are taken at face value, we have not thereby dispensed with the structure of the eliciting event, for now that structure is built into the description of the response itself. Thus in (1'), the elicitor is the presence for o at t of an object which is both red and card-shaped, and the response is not merely moving toward a nearby red object and also toward a card-shaped object (as could be done, e.g., if a juxtaposed black card and red ball were near o at t), but to an object which is *both* red and card-shaped. An even more complicated structure, borrowed intact from the eliciting event, is incorporated into the definite description by which the direction of o's response is given in (2').

While the analysis of these examples could be spun out at greater length, their significance should by now be apparent enough: Many of the empirical generalizations we draw from the most commonplace behavioral data, when written in their simplest and most natural form, do *not* correspond to the mere-presence-of-stimulus-evokes-fixed-response model of elicitation, but make use of the whole eliciting event (i.e., its ingredients *and* its structure) to predict a response which varies from one elicitation to the next. To be sure, these generalizations may perhaps be recast into orthodox S-R jargon, which prefers to speak of a response *as though* it were evoked by a stimulus rather than by a fact involving the stimulus, by conceiving certain logically complex predicates as descriptions of stimuli, and embedding the structure of the eliciting event in the description of the response itself, but by such artificial maneuvers we lose sight of the basic simplicity of the relationship.

It has been customary to think that flexible responding to situations-as-wholes is something which emerges at the level of the higher organisms in complex situations, to the consternation of S-R theorists and the delight of those who are opposed to systemic rigor. Actually, many of the behavioral relationships with which the rigorous behaviorist feels most at home are of this kind, but recognition of this has been thoroughly obscured by our conventional manner of speaking about the external causes of behavior. Once we *ungrudgingly* admit that stimuli cannot in themselves evoke behavior, and adjust the conceptual framework of behavioristics to bring out clearly the logical relations between eliciting event and response event, we will surely find that conditioning data and what we know about the “higher mental processes,” which now appear to be unrelated if not downright inconsistent, are but different instances, and perhaps not even very different instances, of a common set of rigorously formulable behavioral principles.

The Behavioral Analysis of Semantics⁶

Finally, I would like to show how recognition that facts, rather than stimuli, elicit behavior grants important new insight into that most disgracefully neglected problem, the *semantical* dimension of language. Despite the hundreds of works which have appeared in the psychological literature on various aspects of language, one can count on one’s fingers those which have attempted seriously to analyze its nature as a *representational* process. It is only because words *refer to*, or *denote*, or *stand for*, or *signify*, or *designate*, or *are about* other things that they are able to serve as vehicles for thought and knowledge. Yet the nature of the semantic relation between words and things (for which the italicized expressions are, roughly, synonyms) is poorly understood, and for want of an adequate behavioral analysis of Meaning and Reference—for it should be obvious that symbols have whatever cognitive properties they do have largely because of the way people behave with respect to them—a great deal of nonsense has been written by philosophers on matters of ontology and epistemology.

It has long been recognized that the supreme virtue of symbols’ semantical properties is that by proper exploitation of words, we are able to adjust efficiently to conditions of the environment in the absence of direct sensory access to them. The descriptive, declarative, or “cognitive” use of language, as practiced, e.g., by adult humans in the communication of scientific data and theory, is the most powerful adaptive mechanism ever devised by nature; and what is so adapted to upon encounter with a linguistic expression is not (primarily) the fact that a certain sequence of visual or auditory configurations is present, but some other factive possibility signified by this expression. It is thus evident that the representational

⁶At the request of the original editor to shorten the paper, the following closing remarks were omitted from the version published in *Philosophy of Science*.

relation of words to things must somehow consist in the behavioral effects induced by the former being “appropriate” to the latter. However, the trick is to spell out the details of this “appropriateness,” and that task has proved to be surprisingly refractory.

Since the meaning units of a language are, roughly speaking, its individual words, the logical starting point for behavioral analysis of semantics might seem to be determination of what behavioral state an organism needs to be in if a certain word is to designate a certain entity for that organism—e.g., for ‘Tom’ to signify a particular person, or for ‘red’ to refer to the color red. The simplest suggestion that comes to mind in orthodox S-R terms is that a stimulus S_1 , signifies another stimulus S_2 if S_1 , through a process of learning, has come to evoke the same response as S_2 ; and while this hypothesis is quickly shown to be unsatisfactory in its most elementary form, some variation of it has been the foundation of all behavioral analyses of sign processes to date. The latest and most behaviorally sophisticated of these is Osgood’s: “Words represent things because they produce some replica of the actual behavior toward these things . . . A pattern of stimulation which is not the object is a sign of the object if it evokes in an organism a mediating reaction, this (a) being some fractional part of the total behavior elicited by the object and (b) producing distinctive self-stimulation that mediates responses which would not occur without the previous association of non-object and object patterns of stimulation.” (Osgood, 1953)

It is here unnecessary to submit Osgood’s treatment to a detailed criticism, for all theories which seek to interpret symbolic reference as a relationship between the behavior evoked by a symbol and that evoked by its referent are open to a common fundamental objection: Disregarding the behavioral import which occurrence of a symbol has independently of its symbolic role (the flinch elicited by overly loud speech, the effect on a proofreader of a misspelled word, etc.), *the individual words of descriptive discourse, taken in isolation, do not appear to arouse any particular behavioral tendencies, either actual or dispositional*. When language is used in its referential capacity, its smallest division which mobilizes a behavior adjustment is the *sentence*,⁷ or occasionally briefer expressions which are not strictly sentences grammatically but have acquired the force of a sentence.⁸ Thus, the reader will be hard pressed to discover any behavioral tendencies induced in him by hearing

⁷Or more precisely, the *believed* sentence, since a sentence’s behavioral influence is determined not only by its meaning, but also by the degree to which it is accepted. It is not unlikely that a sentence, when “understood” by a listener, predisposes certain behavioral tendencies which are then to some degree held in check or actively inhibited by a mechanism whose effect is subjectively experienced as the extent to which the sentence is believed or disbelieved.

⁸For example, the descriptive components of such exclamations as “Fire!” and “Come!” are rendered by sentences such as “There is a fire nearby, and something must be done about it,” and “I want you to come here.” The analysis of hortatives is complex, and, since their primary function is not a cognitive one, not of major semantical concern.

or seeing the word 'red' in linguistic isolation, except perhaps the tendency to ask, "Well, what about it?". In contrast, one can easily detect sets and perhaps even overt responses produced by hearing (and believing) sentences such as "You just drove through a red light," "A person who fixates long enough on a red object will experience a blue-green afterimage," "John wore a bright red hunting jacket to the funeral yesterday," or "The City Council just voted to raze our red light district." In each of these, the word 'red' contributes to the total effect of the sentence, and in each case we would say that the word 'red' signifies the color red; yet it is hard to think of a common behavior tendency produced by the various occurrences of 'red' when these sentences are presented to someone proficient in English. Obscurantistically, this can be put by saying that the behavioral force of the stimulus word 'red' is "context dependent." Now in fact, this context dependency of single-word effects is reassuring rather than dismaying, for when one *perceives* that he drove through a red light, that fixation on red objects produces blue-green afterimages, etc., the elicitation impact of redness as a stimulus is also "context dependent," thus enhancing the plausibility that symbolic representation of the color red by the word 'red' consists of some relation between the behavioral significances of 'red' and redness. But this cannot be a similarity between the responses evoked by the attribute redness and the geometric pattern 'red', for the simple reason that neither of these stimuli *work* by eliciting a response. Such a relation is not even logically possible—the closest approximation that could occur would be for the response elicited by the fact that redness is present to be similar to the response elicited by the fact that 'red' is present; and not only is this just not so, neither would it help us to understand how the word 'red' contributes to the behavioral prescriptions of sentences which contain it.

Once one recognizes, however, that facts, not stimuli, elicit behavior, it becomes gratifyingly clear that the unit of language which induces an actual behavioral adjustment, overt or propensitory, *should* be the sentence rather than the individual word. For if the way in which language signifies aspects of the world is by predisposing behavior "appropriate" to the latter, then the behavior-producing unit of language should be that which refers to the behavior-producing unit of the environment—and *sentences* are what refer to *facts*. Consequently, traditional behavioristic analysis of semantical relations is not, perhaps, so implausible after all when it proposes that a language component *L* refers to an aspect *W* of the world when the behavioral adjustments predisposed by *L* are sufficiently similar in certain ways to those predisposed or elicited by *W*; the error has been to presuppose that '*L*' and '*W*' in this formulation range over words and stimuli, respectively, whereas we now see that their respective ranges must be sentences and facts. However, this shift leaves still unexplained the logically more basic connection of individual words to what they stand for. I shall say no more on this point except to comment that the explanation is undoubtedly to be found in some

relationship—not necessarily one of similarity—between the behavioral *import* of the word and what the word signifies; and raising the matter of “behavioral import” returns us full circle to pondering the role of stimuli in the elicitation of behavior.

Although an attribute cannot itself be a cause, it nonetheless has causal *import*, as expressed by statements which describe how the involvement of that attribute in an event influences the latter’s causal ramifications. Thus while stimuli do not themselves elicit behavior, they most assuredly have behavioral import. When behavior is construed to be the net effect of established response tendencies elicited by the various stimuli present, the behavioral import—or more precisely, the *elicitation*-import or *cue*-import—of a stimulus can be represented by associating with each stimulus description a function over organisms and times which coordinates a response description to each $\langle o, t \rangle$. The rule for interpreting this function is that whenever a stimulus S is present for o at t , we predict that o will have a tendency at t to make the response identified as the elicitation-import of S for o at t . However, were we to try to preserve this form of analysis for even such simple phenomena as learning to move toward objects which are both red and card-shaped, or toward the largest red card present, we would have to say that elicitation import for o has been acquired neither by redness nor card-shapedness in itself, but only by the compound attribute being-both-red-and-card-shaped or, even more esoterically, by being-the-largest-red-card-present.

Now, it has already been argued that to interpret every elicitor as a mere presence-of-stimuli kind of event is theoretically gratuitous, distorts the empirical generalizations on which behavioristics is grounded, and is hence with overwhelming likelihood incorrect. It seems much more likely that in elicitation of, e.g., o ’s going toward x at time t by the fact that x is a red card present to o at t , redness and card-shapedness do have elicitation-import for o at t , except that this cannot be represented so simply as by associating with each stimulus a specific response tendency for o at t . That redness and card-shapedness themselves, rather than merely the compound attribute being-both-red-and-card-shaped, acquire elicitation-import when o is trained to go to red cards will be empirically substantiated—if it is—by discovering that this sort of training also has certain effects on the ways in which o now responds to other kinds of situations involving redness or card-shapedness.

In short, I am suggesting that an adequate theory of behavior must represent the elicitation-import of a stimulus S for o at t by something more formally complex than description of a specific response tendency. What form of expression these elicitation-imports will take in the technical language of behavioristics, I have no present idea, except that if f is a fact about the immediate environment of o at t , their rule of interpretation must derive the behavior tendency elicited

by f from o at t from the elicitation-imports for o at t of the attributes component in f . Before any serious formalization can be attempted in this direction, however, we shall need a great deal more empirical data concerning the manner in which conditioning with respect to one kind of event affects responses to events of somewhat different but overlapping composition. And this research, in turn, can be pursued effectively only after we re-examine a number of the empirical generalizations which have already been established about behavior, and write them in logically proper, *unabbreviated* form, so that we can see just what it is that they actually do say.

References

- Adams, D. K. (1954). *Learning and explanation. learning theory, personality theory, and clinical research: The Kentucky Symposium*. New York: Wiley.
- Hilgard, E. R., & Marquis, D. G. (1940). *Conditioning and learning*. New York: Appleton-Century-Crofts.
- Osgood, C. E. (1953). *Method and theory in experimental psychology*. Oxford: Oxford University Press.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Spence, K. W. (1942). The basis of solution by chimpanzees of the intermediate size problem. *Journal of Experimental Psychology*, *31*, 257-271.
- Woodworth, R. S. (1921). *Psychology* (1st ed.). New York: H. Holt & Co.