Joint Control and the Generalization of Selection-Based Verbal Behavior

Barry Lowenkron
California State University, Los Angeles

Although the acquisition of selection-based verbal behavior can be ascribed to the acquisition of a conditional discrimination, such an account cannot explain any generalization of the behavior to novel verbal stimuli. The problem is that printed and spoken words and phrases do not vary on continuous dimensions that would support stimulus generalization. Both conceptual analysis and empirical evidence suggest that an alternate form of stimulus control, joint control, can more readily account for acquisition and generalization of these performances. The fact that joint control depends on topography-based behavior implies that generalized selection-based behavior is not an alternative to topography-based behavior but depends on its prior development.

Skinner (1957) characterizes verbal behavior as behavior by a speaker that is reinforced by those who listen and act in accord with what is said. For verbal behavior to be effective in providing reinforcement to the speaker, the stimuli it produces must be distinctive enough to control the behavior of the listener. Most commonly, these distinctive products arise from a collection of diverse response topographies in the spoken and written repertoire. But there is another possibility: distinctive stimulus products may also result from a single response; as where a person selects different symbols, pictures, or printed words by pointing to them.

TOPOGRAPHY-BASED AND SELECTION-BASED VERBAL BEHAVIOR

In recent articles Michael, (1985) and Sundberg and Sundberg, (1990) have begun to develop and explore the distinctions between topography-based and selection-based verbal behavior. Topography-based verbal behavior is characterized by a correlation between distinct response topographies (i.e., writing and speaking words) and their controlling antecedents. In the topography-based tact, for example, the antecedents are nonverbal stimuli (i.e., objects) and the verbal response topography that each stimulus controls may be said to describe, or name, the stimulus (e.g., saying or writing "book" or "pencil" to the respective objects, tacts each of them).

By contrast, in selection-based verbal behavior, a single selection response, such as pointing, is controlled by two types of stimuli: the set of comparison stimuli comprising an array to be selected from, and a single sample stimulus specifying which comparison may currently be selected for reinforcement. Either or both types of stimuli may be verbal (e.g., a written or spoken phrase) or nonverbal (e.g., an object). Where the sample stimulus is verbal and the comparison stimuli are nonverbal, Michael (1985) describes the selection behavior as *manded stimulus selection.* In such a task, the subject would, for example, select a red square in response to a verbal stimulus: the experimenter's mand to "Find the red square."

Where the sample is nonverbal and the comparisons are verbal stimuli, Michael (1985) terms the response a *selection-based*

Requests for reprints of this article should be addressed to Barry Lowenkron, Dept. of Psychology, California State University, Los Angeles, Los Angeles, CA 90032.

tact. Here the subject would select the printed phrase *red square* when shown a red square as the sample. Finally, where the sample and comparison are both verbal and do not have point-to-point correspondence, the performance is a *selection-based intraverbal* (Sundberg & Sundberg, 1990). Here, the subject would select the printed comparison *red square* in response to the vocal sample "red square."

RESPONSE PROBABILITY AS A DETERMINANT OF COMPARISON SELECTION

In his description of selection-based behavior, Michael (1985) treats the sample and comparison stimuli as elements of a conditional discrimination. Under this form of stimulus control, the probability of a selection response is heightened in the presence of a discriminative stimulus (S^D). However, the capacity of any particular stimulus to function as an S^D is in turn controlled by the status of another stimulus: the conditional stimulus. As applied to the selection task, this analysis proposes that the capacity of comparison stimuli to function as S^Ds is determined by the state of the sample stimulus. As a result, different comparisons may evoke a selection response, depending on the state of the sample stimulus.

According to this account, any two stimuli may be arbitrarily paired by making one stimulus function to control the probability of a selection response to the other. Such an account, however, treats *all* stimuli as if they were arbitrarily paired. The account is thus fundamentally deficient in that it ignores consistent relations between stimuli. For example, the selection of a blue comparison in the presence to a blue sample is treated as equivalent to the selection of a circle comparison in the presence of a blue sample. Both result from the heightened probability of a selection response to a comparison in the presence of the sample. The identity relation between the colors in the first case is thus ignored (Carter & Eckerman, 1975). As a result of this deficiency, the conditional discrimination account cannot explain the occurrence of generalization based on consistent relations such as identity. (See Lowenkron, 1984 for a further discussion of this problem.)

The deficiency remains when the stimuli are verbal. Consider for example, a manded stimulus-selection task in which subjects are trained to select appropriate comparisons when the experimenter gives mands such as: "Select the square above the triangle" and "Select the triangle below the circle." If each mand merely functions as a sample, making the evocation of a selection response in the presence of a particular comparison stimulus more probable, then the prepositional relations *above* and *below* do not enter into the control of behavior during training, and there is no basis for generalization to novel phrases such as "Select the square below the circle."

A similar case prevails with logical relations. Training subjects to select particular objects in response to mands for objects *green* and *yellow* and *blue* or *red* provides no basis for generalization to novel mands for objects *blue* and *red*, and *green* or *yellow*.

Where the comparison stimuli are verbal (e.g., printed words), an account phrased strictly in terms of response probability faces additional problems because this account only requires that printed-word comparisons *evoke* a comparison selection response (e.g., pointing). The role of a textual response to the comparison stimulus is not considered. But to assume that printed-word comparisons are responded to only on the basis of their physical characteristics, and are not read before they are selected, is simply implausible. It also precludes any account of generalized performance.

Consider for example the selection-based intraverbal: the subject is given a spoken word and must select a printed word. One cannot appeal to simple stimulus generalization along the physical dimensions of the stimuli to explain why, when given "receive" as a novel sample, the subject will select the printed *receeve* rather than *deceive*. Both differ from the correct spelling by one letter, why does the subject pick *receeve*?

Similar problems exist when selection-based intraverbals appear in the repertoire with no history of reinforced selection. Thus, despite the absence of prior reinforcement for the behavior, all readers can select the printed *Schwartzkopf* in response to a spoken sample. Likewise, the printed word *kauphy* will be selected in response to the spoken sample "coffee" despite the fact that the former shares no common letters with the familiar spelling and should therefore be less likely to evoke the selection response than the printed word *coffin*.

The same problems arise in accounting for the generalization of selection-based tacts. With nonverbal samples, there is no regular correspondence between the visual characteristics of printed comparisons and the samples they describe. The printed phrases *line over box* and *box under line*, although visually distinct, are tacts for the same nonverbal sample. The phrase *box over line* is visually more similar to the latter even while the sample it describes is visually distinct. Clearly, learning to select appropriate phrases on the basis of their physical characteristics provides no straightforward basis for generalization.

A more plausible explanation of these kinds of behavior would recognize that subjects read the comparison words before selecting them, and that this textual behavior causes the sample and the correct comparison to control identical pronunciations. But an account based on this identity relation would still not explain generalization to novel stimuli. To say subjects select novel comparisons that evoke a pronunciation identical to the sample merely shifts the burden of explanation to a cognitive mediator: to a *concept of identity* with regard to pronunciations.

JOINT CONTROL AND COMPARISON SELECTION

There is, however, a simple remedy which does not involve a non-behavioral process. The role of a concept of identity can be replaced by the operation of a pair of verbal relations (Lowenkron 1988, 1989). These verbal relations jointly exert control over the topography of a single verbal response that precedes, and thus mediates, the actual selection response.

Consider, for example, a task in which the subject is given a spoken-word sample and must select the corresponding printed word from an array. Ostensibly, this is a selection-based intraverbal. After minimal training, a normal reader will perform errorlessly: even pointing to words and phrases (i.e., *Saddam Hussein*) for which a selection response was never previously reinforced.

Clearly, in this task the subject must read each of the comparison words in the array until one is encountered which allows the subject to emit the same response topography by reading (i.e., as a textual) as is emitted as an echoic of the spoken-word sample. Thus, as the subject moves through the printed comparisons, attempting to repeat the sample phrase as an echoic in the presence of each comparison while responding to the comparison as a textual, most of the comparisons will control textuals with *conflicting* topographies. Only the correct comparison will evoke a response topography which is *both* an echoic with respect to the sample, (that is, a repetition of the sample) and a textual with respect to the comparison (that is, an accurate reading of the comparison). The topography of the repetition is thus under *joint control* of the sample (as an echoic or self-echoic) and the correct comparison (as a textual).

The emission of this primary operant (i.e., the repetition) is then the occasion for a selection response such as pointing. This selection response is itself verbal behavior. Because it is only emitted when the prior response occurs under joint control, and because it informs the listener (in this case the experimenter) of this control, the selection response is a descriptive autoclitic (Skinner, 1957, chap. 12). It tacts the unique condition of stimulus control wherein the echoic and textual control a common topography. And so, in this account recognition of the "sameness" of the response topographies evoked by a sample and by the correct comparison is not treated as a cognitive event on the part of the subject, rather it is a feature of stimulus control over an autoclitic.

JOINT CONTROL AND GENERALIZATION

Joint stimulus control fosters generalization of selection-based behavior because it provides a common antecedent, (i.e., the emission of any response topography under joint control) for a common comparison-selection response (i.e., pointing). Thus, any sample-comparison combination that evokes joint control of a response topography, will be so reported by the autoclitic selection response.

The role of joint control in a generalized performance was demonstrated by Lowenkron (1988). To permit the direct measurement of each component of the performance, overt mediating responses were trained. The task in this study was generalized delayed identity matching to sample. Retarded children were explicitly trained in four component responses: (a) use a handsign to tact the current sample shape; (b) maintain the current handsign, unchanged, over a delay interval; (c) when they appear, correctly tact one of the comparison stimuli without changing the current handsign (joint control); and (d) push on the comparison that allows joint control in order to select it. Since the only comparison that met all of these conditions was the one with the same shape as the sample, a correct performance of all the components necessarily generated an identity match.

The handsigns, of course, were a variety of verbal behavior. Requiring that the sample handsign remain unchanged through out a trial provided a performance functionally equivalent to a self-echoic (Skinner, 1957). Coding each shape with a different handsign required a topography-based tact, and thus only the correct comparison allowed a particular handsign to occur under joint echoic-tact control. This control was then reported by the subject pushing on the comparison—a descriptive autoclitic for joint control.

When the subjects were subsequently tested for generalization of the performance to novel shapes, none was detected —even after the subjects received discrimination training with the novel stimuli. However, as soon as the subjects learned handsigns to tact the novel shapes, generalized matching appeared immediately and from the very first test trial. Presumably, the matching performance generalized at this point because the novel stimuli could now enter into joint control of the handsigns and were so reported by the selection response.

Other studies (Lowenkron, 1988, 1989) have replicated these findings, but with abstract relations such as relative size (greater or lesser), distance, and spatial orientation. In both studies, once the performance was trained under joint control, generalized delayed matching again depended solely on the availability of topography-based responses with which to tact the novel stimuli.

Strictly speaking, these tasks, based as they were on nonverbal stimuli (i.e., geometric shapes), might be classified as nonverbal selection-based tasks. But to the extent that they were mediated by verbal behavior, such a distinction seems superfluous. Indeed, Lowenkron (1991) replicated the findings reported in these studies using vocal mediating responses and a manded stimulus selection task.

Taken together, these studies suggest that other generalized selection-based performances, in particular, selection-based verbal behavior, may depend on joint control[1]. Conceptual analysis bears this out. For example, in the case of the selection-based intraverbal, given the spoken sample "receive," the printed comparison *receeve*, but not *deceive*, allows the subject to repeat the pronunciation as both an echoic of the sample, and as a textual. This is joint echoic-textual control. Similarly, in the presence of the spoken *Schwartzkopf*, the printed version would be selected, because it alone permits joint control over the repetition of the spoken sample. Likewise, it is

[1]All of this is not to say that matching may never be acquired without mediating responses. The focus here, however, is on the conditions under which matching generalizes.

kauphy, but not *coffin*, that allows a repetition of the spoken sample "coffee" under joint echoic-textual control. If the subject has previously learned to select comparisons that permit joint control, then generalization in these examples should be immediate.

The case of the selection-based tact is similar. Here, the sample is a nonverbal stimulus, and so it is the subject that provides the verbal sample by making a topography-based tact to the sample object. Thus, when the box and line are given in a new relation as a novel sample, generalization under joint control depends only on the subject's ability to tact this sample. Doing so converts it into a verbal stimulus, and the remainder of the performance becomes a selection-based intraverbal. The subject now must repeat the novel sample tact (i.e., "Box to left of line") as a self-echoic and select the printed word comparison that controls the same topography under textual control.

Finally, in manded stimulus selection, where the sample is verbal but the comparisons are not, joint control requires the subject to tact the comparisons. Now reinforcement is contingent upon selecting the comparison permitting joint echoic-*tact* control. Thus, given the experimenter's mand to "Find the red square," the subject must select the comparison that allows the topography "red square" to be emitted as both an echoic of the experimenter's mand, and as an accurate tact of the stimulus object.

THE PRIMACY OF TOPOGRAPHY-BASED BEHAVIOR

Clearly, joint control depends on topography-based behavior. Therefore, to the extent that generalization of selection-based behavior depends on joint control, it ultimately depends on topography-based behavior. Generalized selection-based verbal behavior is thus dependent on, rather than an alternative to, topography-based verbal behavior.

Aside from the studies cited earlier, there is other evidence for this dependence. In a recent study Sundberg and Sundberg (1990) made a direct comparison between the two kinds of behavior. Some subjects were trained in topography-based behavior. They learned to emit a common handsign; first as a topography-based tact for a particular sample object, and then as a topography-based intraverbal for a spoken sample word. Other subjects were trained in selection-based behavior. They responded first to a sample object and then to a spoken word by selecting a common symbol; thereby acquiring respectively a selection-based tact and an intraverbal. All subjects learned their respective types of common responses for each of three object-word pairs.

Overall, the common selection-based responses were learned far more slowly than the topography-based responses. This disparity suggests that subjects had more to learn in the former condition: perhaps, mediating-response topographies to control selection-based behavior.

The emergence of manded stimulus selection within each word-object pair was then measured on unreinforced test trials interspersed among reinforced intraverbal and tact trials. On these unreinforced test trials one of the trained words was spoken by the experimenter, and the subject had to select the object that shared the common handsign (for subjects trained in the topography-based behavior) or that shared the common symbol (for subjects trained in the selection-based behavior). What is critical here is that over trials appropriate selections of the stimuli manded by the experimenter only developed in subjects trained with the topography-based behavior.

Subsequently, those subjects trained in the topography-based behavior were trained with three new word-object pairs in the selection-based behavior and then retested for the emergence of manded stimulus selection. It did not appear. Thus, even with a history of accurate manded stimulus selection, subjects could not reproduce the behavior if topography-based responses were unavailable. Selection-based responses, although also common to the word and object to be matched, were inadequate to produce the behavior. What seems critical is not just the availability of a common response, but a common response topography. Manded stimulus selection, a selection-based verbal behavior thus depended on mediation by a topography-based behavior.

Finally, it should be noted, many features of the procedure implicate joint control. If the subject responded to the spoken word with the handsign learned during intraverbal training, then the correct object was always the one for which the same handsign had been trained as a tact. With both of these responses being reinforced on the interspersed training trials, it is not surprising that manded stimulus selection emerged on the unreinforced test trials. All of the components necessary for joint control were being maintained during the test.

SUMMARY

Both conceptual analysis and experimental data suggest that the flexible, generalized, selection of stimuli in response to their description cannot be described by the automatic compounding of selection response probabilities functioning within a conditional discrimination. Rather, it suggests that the selection of stimuli in response to their specification, or to the specification of relations between stimuli, requires precurrent verbal behavior: the generation of echoics, textuals, tacts and autoclitics. Replacing the automatic process of selection-response evocation with these events provides a behavioral account of what it is that subjects so frequently do before they respond: it explicitly describes thinking as precurrent verbal behavior.

REFERENCES

Carter, D.E., & Eckerman, D.A., (1975). Symbolic matching by pigeons: Rate of learning complex discriminations predicted from simple discriminations. *Science, 187*, 662-664.

Lowenkron, B. (1984). Coding responses and the generalization of matching-to-sample in children. *Journal of the Experimental Analysis of Behavior, 42*, 1-18.

Lowenkron, B. (1988). Generalization of delayed identity matching in retarded children. *Journal of the Experimental Analysis of Behavior, 50*, 163-172.

Lowenkron, B. (1989). Instructional control of generalized relational matching to sample in children. *Journal of the Experimental Analysis of Behavior, 52*, 293-309.

Lowenkron, B. (1991). *Vocal mediation of generalized matching to sample in children*. Manuscript submitted for publication.

Michael, J. (1985). Two kinds of verbal behavior plus a possible third. *The Analysis of Verbal Behavior, 3*, 2-5.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton Century-Crofts.

Sundberg, C. T., & Sundberg, M. L., (1990). Comparing topography-based verbal behavior with stimulus-selection-based verbal behavior. *The Analysis of Verbal Behavior, 8*, 31-41.