JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1998, **69,** 327–354 NUMBER 3 (MAY)

SOME LOGICAL FUNCTIONS OF JOINT CONTROL

BARRY LOWENKRON

CALIFORNIA STATE UNIVERSITY, LOS ANGELES

Constructing a behavioral account of the language-related performances that characterize responding to logical and symbolic relations between stimuli is commonly viewed as a problem for the area of stimulus control. In response to this problem, the notion of joint control is presented here, and its ability to provide an interpretative account of these kinds of performances is explored. Joint control occurs when the currently rehearsed topography of a verbal operant, as evoked by one stimulus, is simultaneously evoked by another stimulus. This event, the onset of joint stimulus control by two stimuli over a common response topography, then sets the occasion for a response appropriate to this special relation between the stimuli. Although the mechanism described is simple, it seems to have broad explanatory properties. In what follows, these properties are applied to provide a behavioral interpretation of two sorts of fundamental, putatively cognitive, performances: those based on logical relations and those based on semantic relations. The first includes responding to generalized conceptual relations such as identity, order, relative size, distance, and orientation. The second includes responding to relations usually ascribed to word meaning. These include relations between words and objects, the specification of objects by words, name-object bidirectionality, and the recognition of objects from their description. Finally, as a preview of some further possibilities, the role of joint control in goal-oriented behavior is considered briefly.

Key words: cognition, generalization, verbal behavior, joint control, stimulus control, humans

> Our knowledge of another person is limited by accessibility, not by the nature of the facts. We cannot know all there is to know, as we cannot know all we should like to know about the worlds of physics and biology, but that does not mean that what remains unknown is of a different nature. As in other sciences, we often lack the information necessary for prediction and control, and must be satisfied with interpretation, but our interpretations will have the support of the prediction and control which have been possible under other conditions. (Skinner, 1974, p. 176)

If a child is shown a set of red and blue squares and triangles and is then trained to select from this set the colored shape that is identical to a designated sample, it would not be surprising to find that this identity-matching performance generalizes, with no additional training, to green and yellow blocks shaped as diamonds and circles. And similarly, if a child is first trained to select these red and blue blocks in response to spoken requests specifying their color and shape, and is then trained to name the colors *green* and *yellow* and the shapes *diamond* and *circle*, it would be equally unsurprising to find that with no further training the child showed generalized selection: now picking untrained combinations of blue, red, green, and yellow squares, triangles, diamonds, and circles in response to their spoken names.

Simple as they are, these generalized logical and semantic performances exemplify classes of behavioral phenomena that are vitally important to an understanding of human functioning. They are important, first of all, because they involve responding on the basis of relations that exist between stimuli, rather than on the basis of the physical features of the stimuli themselves. They thus represent a critical advance in the logical and linguistic behavior of the individual from the concrete to the abstract. These performances are also important because, once trained, they allow novel behavior to appear without additional training, and so they represent highly efficient forms of behavior: Thus, once a generalized identity-matching performance has been acquired, it may appear with an infinitely large set of objects and events. Likewise, in the normal course of language development, words ultimately come to be emitted and responded to in an unlimited va-

Preparation of this article was supported in part by Grant HD 29232 from the National Institute of Child Health and Human Development.

Correspondence regarding this article should be addressed to Barry Lowenkron, Department of Psychology, California State University, Los Angeles, Los Angeles California 90032 (E-mail: zlowenk@calstatela.edu).

riety of untrained combinations and contexts. Finally, as will become evident, these performances are important because they combine with other behavior to form a seemingly endless variety of novel performances with new, as well as with familiar, stimuli.

Despite its apparent importance, however, a parsimonious, empirically supported, operant account of *all* of these sorts of performances has not been forthcoming. The nature of the problem for a behavioral account may be illustrated by considering a conditional discrimination task in which the stimuli are arbitrarily paired; for example, where reinforcement for selecting particular comparison shapes is contingent upon the color of the sample stimuli. In such a task the sample color is said to function as a *conditional stimulus* that acts to determine which of the comparison shapes will serve as the discriminative stimulus (S^D) for a selection response (such as pointing). It is understood that this selection response arises from the history of reinforcement for such selections in the presence of these stimuli. Henceforth, we may describe this form of stimulus control over the selection response as *unmediated stimulus selection.*

For the radical behaviorist, the unmediated selection account has a unique strength in that it is nonrepresentational: Following a history of reinforced selections, one is not said to hold a memory or other representation of the stimulus being sought for selection; rather, what is retained is the changed likelihood of a selection response to a particular comparison in the presence of a particular sample (Skinner, 1969, pp. 273–274).

With this strength, however, there come some serious limitations in describing the emergence of untrained responding based on relations between stimuli; this is because the unmediated selection account does not recognize such relations, but rather treats all stimuli as if they were *arbitrarily paired.* Thus, on this account, there is in principle no difference between selecting a *square* comparison in response to a *blue* sample (arbitrary matching) and selecting a *blue* comparison in response to a *blue* sample (identity matching). In both cases the sample is said to cause a comparison to function as an S^D for the selection response. But by describing performance in an identity-matching task as if it were arbitrary matching, this approach does not recognize, nor in any way consider, the crucial element of identity matching; namely, the sharing of common features by the sample stimulus and the matching comparison. As a result, such an account is intrinsically unable to describe a performance, such as generalized identity matching, that is based on this formal commonality of stimulus features.

The unmediated stimulus selection account encounters similar problems in explaining the emergence of untrained relations between words and objects. Thus, if this account were applied to describe a task in which reinforcement was contingent upon the selection of objects in response to their spoken names, the names would be described as conditional stimuli that caused the named objects to function as S^Ds for the selection response. But such an account could not explain how subsequent practice by the subjects in emitting novel names in response to novel objects would produce accurate selection with these stimuli in response to the names when spoken by another. Here, accurate selection would appear with none of the reinforcement history necessary for the novel stimuli to function as conditional and discriminative stimuli.

Together, these examples suggest that an operant account that describes relations between stimuli solely in terms of the control individual stimuli exert over a selection response is simply too sparse to describe the many abstract performances that humans emit and that rely on relations between stimuli. Other authors (Schoenfeld & Cumming, 1963, p. 241; Skinner, 1969, chap. 6) have reached similar conclusions, appreciating the role mediating responses play in abstract behavior. The remainder of this paper is an attempt to realize their considerations by describing a simple and parsimonious account of some logical and semantic performances based on the role verbal behavior plays in mediating other, directly observable behavior.

The Nature of Joint Control

The present account, although it preserves the ordinary definition of stimulus control, does so in a manner that appreciates stimulus features and relations between stimuli bearing such features. It can do so because it is based on the effect of two S^Ds acting jointly

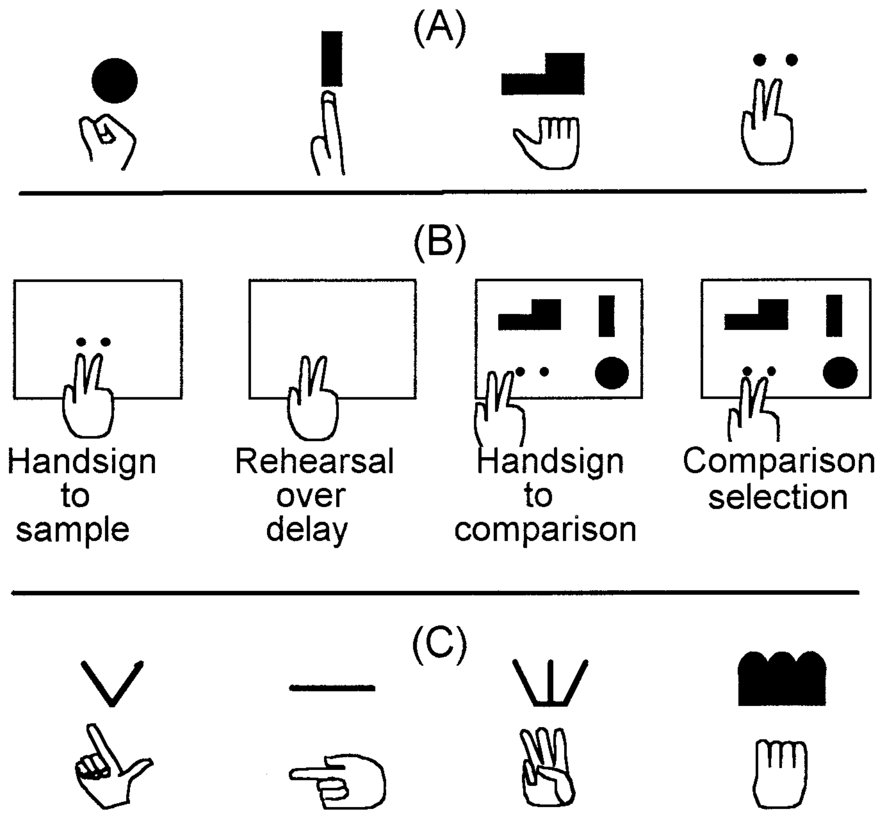

Fig. 1. A generalized identity matching-to-sample performance mediated by hand signs. (A) The training-set stimuli and their associated hand signs. (B) The four component responses of the identity-matching performance: Subjects must make the appropriate hand sign to the sample when it appears, maintain the hand sign over the delay interval, and then, while rehearsing it, attempt to make the hand sign to one of the comparisons. Finally, they select the comparison to which they make the rehearsed hand sign. (C) The transfer-set stimuli and their associated hand signs.

to exert stimulus control over a common response topography. Thus, the account requires no novel concepts; ultimately, stimulus control remains the product of differentially reinforced responding in the presence of different stimuli, and consequently, although mediated by joint stimulus control, so does stimulus selection.

Examples of Joint Control

The nature of joint control may be illustrated by a study in which nonverbal retarded children were trained to use hand signs to name stimuli in a generalized identity-matching performance (Lowenkron, 1988). The subjects first acquired hand signs for the four shapes of the training set illustrated in Panel A of Figure 1. They were then trained to use these signs to mediate an identity-matching performance (Panel B). As a result of this training, subjects learned to make the sign appropriate to whatever shape comprised the current sample and then maintain the sign over the delay interval, leaving it unchanged as the comparisons were presented. The children also learned to make the still-unchanged hand sign to that comparison for which it was appropriate. When all of these components were performed correctly, this necessarily produced an identity match.

Immediately after this performance was trained, the occurrence of generalized identity matching was assessed with the transfer-set shapes shown in Panel C, but it did not appear. Then, in the next phase, each of the hand signs shown in Panel C was trained to the corresponding transfer-set shape. In the subsequent retest, generalized identity match-

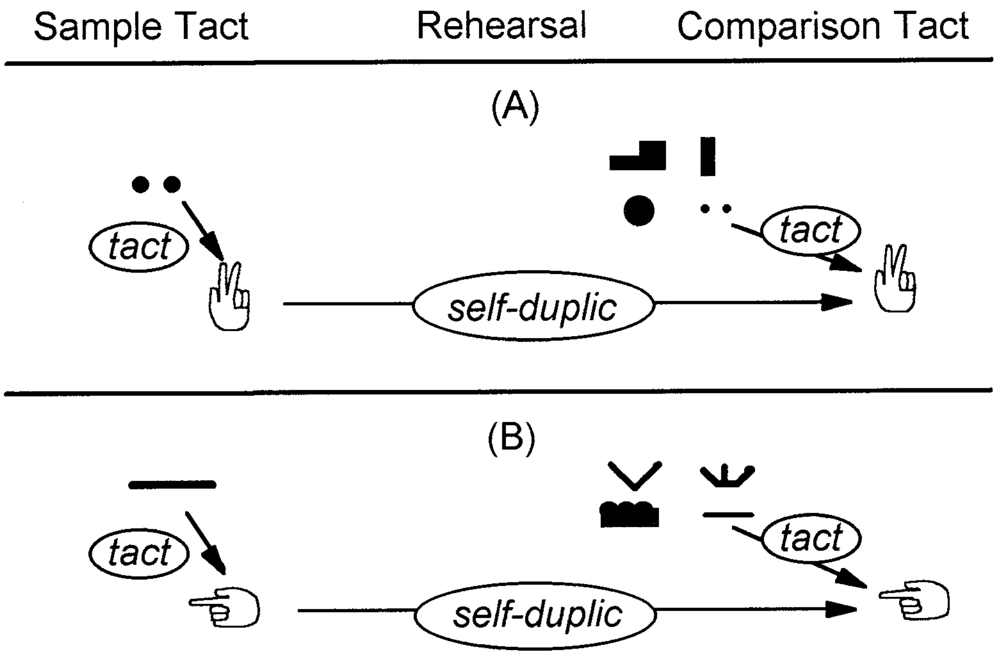

Fig. 2. The basis of generalized identity matching. The particular characteristics of the stimuli are irrelevant because the same verbal relations, tact and self-duplic, are evoked by any comparison that is identical to the sample. Thus, to select the comparison that is identical to the sample, the subject need only select whatever comparison evokes the hand sign currently being rehearsed.

ing appeared immediately. Thus, once subjects acquired hand signs (i.e., names) to the novel shapes, the identity-matching performance immediately generalized to these stimuli.

Why this pattern of behavior produces generalized matching can be simply explained in terms of the combination of verbal relations between hand signs and stimuli illustrated in Figure 2. This figure shows identity matching with stimuli from the training set (Panel A) and the transfer set (Panel B). In both cases, the hand signs to the sample shapes, as topographies correlated with nonverbal stimuli, may be described as the verbal operant Skinner (1957) called a *tact*.[1] Rehearsing the hand sign through the delay interval and during the comparison presentation is a kind of *self-duplic* repetition of the sample tact (Michael, 1982), and finally, making the hand sign to a comparison is again a tact. In each of these cases the correct comparison stimulus is unique in that it is the only comparison that allows the subject to make one hand sign that is both a self-duplic (i.e., a repetition) with respect to the sample tact and also (i.e., jointly) a tact for that comparison. Thus, here, generalized identity matching with novel stimuli only requires that on each trial the subject select the comparison that allows rehearsal of the current duplic topography to continue, but now under an additional source of stimulus control.

[1] The verbal operants mentioned throughout this paper are as follows. A *tact* is a verbal response whose features (spoken or signed) are under stimulus control of a nonverbal object or event or its properties. Names and descriptions of stimuli emitted by a speaker are tacts. An *intraverbal* is a verbal response whose features are under stimulus control of a verbal stimulus, as in a memorized sequence wherein each word evokes emission of the succeeding word. A *textual* is a response whose features are under the control of the features of a printed verbal stimulus (e.g., reading) such that there is a consistent correspondence between the features of what is printed and the features of what is emitted. *Duplics* are a class of verbal operants that are characterized not only by a correspondence between the features of the controlling stimuli and features of the resultant responses but by their formal similarity as well. If the behavior imitated is vocal, then the duplic is called an *echoic.* Imitating a handsign would be a *mimetic.* Repeating or rehearsing one's own behavior would be a *self-duplic,* either *self-echoic* or *self-mimetic.*

To illustrate the process with more conventional behavior, consider what happens when one is given a printed six-digit number as a sample and is asked to find the same number from among a page of such numbers randomly ordered. While searching, one may vocally rehearse the sought-after number. When an entry is encountered that allows the rehearsed numbers to be entirely emitted as a series of tacts for the numbers in that entry, searching ceases; for at that moment the subject is saying the numbers under the joint control of both his rehearsal and also the printed numbers now encountered. As with the hand signs, there is a change in the stimuli currently evoking the rehearsed topography. Here it is a change from self-echoic to joint self-echoic/tact control, and it is this change that in turn evokes a selection response such as pointing to the specified number.

In essence, the onset of joint control is a stimulus event that arises with the appearance of a second source of control over a rehearsed topography. It is, we may speculate, the feeling that arises from saying, reading, or doing something in response to a new stimulus that has previously been (or is currently being) practiced as a rehearsal.

A dramatic example of this sort of control can be experienced by first reading the puzzling phrase *Mairzy doats an doesy doats an dlidl lamsy divy*. This phrase, from an old popular song, may not be intelligible to many. For these people there will be a difference between their first experience reading the phrase and their experience rereading it after having read the clarifying phrase: *Mares eat oats and does eat oats, and little lambs eat ivy*. The difference is that in rereading the puzzling phrase, the words are emitted under joint control, both as a repetition of the clarifying phrase and jointly as textual responses to the novel spellings of the puzzling phrase. And so, if the puzzling phrase had been one of several comparisons to be selected among in response to the sample *Mares eat oats* . . ., only the onset of joint control would distinguish the correct comparison from other, similar phrases such as *Mairy doats* . . . and *Maizy oats* . . . and thereby set the occasion for its selection.

The Mechanism of Joint Control

To describe the mechanism of joint control more explicitly, it is useful to first distinguish this sort of control from the cases Skinner describes in which several variables control a single verbal response. Thus, in the case of *multiple causation* (Skinner, 1957, p. 228), several stimuli contribute strength to the initial evocation of a single response, and under *supplementary stimulation* (p. 254), one or more stimuli contribute strength to the ultimate evocation of a topography that has been determined by prior events. Crucial here is the fact that in both of these cases multiple variables contribute directly to the evocation of the response of interest.

In contrast, with joint control, the response of interest, the selection response, is not directly evoked by multiple variables that contribute to its response strength. Rather, as we have seen, the selection response is evoked by a single event: the occurrence of joint control over some other topography, generally the topography rehearsed as the duplic. Thus, when the appropriate comparison was encountered, the rehearsed hand sign was no longer solely a self-duplic, because now, with no change in its topography, it could also be emitted as a tact for that shape. Likewise, when the appropriate six-digit number was encountered, the rehearsed sequence could now be emitted as a tact. In both of these cases, and more generally, it is this onset of a second, already acquired, source of stimulus control over a rehearsed topography that in turn evokes the selection response.

Joint control is thus an event that is independent of any particular stimulus but is specific to the relation between stimuli. It is this feature that permits joint control to serve as the basis for generalized responding. Thus, as Figures 1 and 2 illustrate, although joint control was specific in that it could only occur with one comparison for each sample, the joint control event itself, the onset of a second source of stimulus control over an already rehearsed topography, was common to all occasions in which a comparison was encountered that allowed the currently rehearsed hand sign to also be emitted as a tact. It was this generic event, the onset of joint control over the rehearsed operant, that acted as the antecedent (i.e., as the S^D) to the actual comparison-selection response and thereby enabled generalized matching.

Given all of the above, it is clear that under joint control the comparison-selection response itself has several noteworthy properties. First, unlike the case of unmediated selection, where it is the individual stimuli that evoke selection responses, under joint control a selection response does not report about (i.e., is not controlled by) the stimuli themselves (verbal responses to the samples and comparisons do this), nor does the selection response directly report about any formal relation between the stimuli. Rather, the selection response reports about (i.e., is occasioned by) the onset of an additional verbal relation over a common, already emitted topography. Thus, as illustrated in Figure 3, the pointing response reports that the current hand sign topography, already occurring as a self-duplic, now also occurs as a tact evoked by the comparison being pointed to. In Skinner's terminology the selection response is a descriptive autoclitic: a tact controlled by the events that control other verbal behavior (Lowenkron, 1991; Skinner, 1957 pp. 313–314). In the present case the selection response tacts the onset of an additional source of control over a rehearsed verbal topography. That this autoclitic is a selection-based (e.g., pointing) rather than a topography-based (e.g., saying a word) tact is irrelevant (Lowenkron, 1991; Michael, 1985): One might as well name the stimulus that evokes joint control as point to it.[2]

And so, although it is true that the hand signs themselves may be described as coding responses (Cumming, Berryman, & Cohen, 1965) or as differential sample responses (Urcuioli, 1985), I believe it is highly significant that the vocabulary that describes the relations that comprise joint control comes, unmodified, from Skinner's description of verbal behavior (Skinner, 1957). The many complex and abstract patterns of responding described throughout this account, normally ascribed to semantic, symbolic, or logical relations, as well as to purpose and goal orientation, consist of nothing more than the combination of Skinner's verbal operants. It is thus a perfectly behavioral instantiation[3] of the relation between thought and language (Vygotsky, 1962).

A Definition of Joint Control

From the foregoing, a complete definition of joint control can be derived: Joint control is a discrete event, a change in stimulus control that occurs when a response topography, evoked by one stimulus (e.g., the sample) and preserved by rehearsal, is emitted under the additional (and thus joint) control of a second stimulus (e.g., the comparison).

As verbal behavior, rehearsal will be some form of self-duplic such as a self-echoic, whereas control by the comparison stimulus may be exerted as a tact, a textual, or some other verbal operant, depending on the nature of that stimulus. As with any source of stimulus control, joint control may be inferred, but it is not directly observable in others. But, as with other verbal operants, it is an event that may be accessible to the speaker and may be reported by a variety of operant (autoclitic) responses including, but not limited to, selecting a stimulus. The notion of joint control is also a parsimonious concept derived from ordinary sources of stimulus control. It merely involves a compounding of two existing sources of stimulus control over the occurrence of a single topography. Its uniqueness lies in the way this compounding functions, not in the incorporation of some newly invented mechanism: logical, behavioral, or cognitive.

Finally, it should be pointed out that in what follows the form of the autoclitic response to joint control is not always specified. In these cases it is safe to assume, and the context should make clear, that the response is an autoclitic selection response of the sort illustrated in Figure 3. Other cases will be described as they arise.

---

[2] Some topography-based autoclitics reporting other sources of control would be "I see . . . ," which reports that what follows is a tact; "You said . . . ," which reports echoic control; and "That reminds me . . . ," which reports an intraverbal source of control. Joint control assumes only that an autoclitic can report about (be controlled by) two sources acting simultaneously.

[3] S^D control, beyond any doubt, was a feature of behavior well before the evolution of mammals. It has thus had ample time to become a "hard-wired" aspect of our anatomy and our behavior. Joint control, on the other hand, could evolve only *after* the acquisition of verbal behavior. It was thus acquired very recently, and as a result, it is implemented in our repertoire as software (i.e., it is simulated through a concatenation of simple operants).

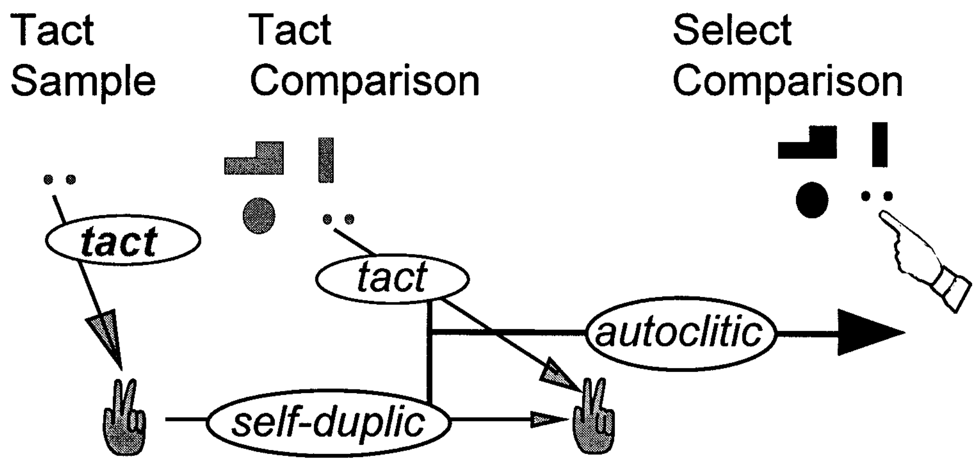

Fig. 3. The selection response as an autoclitic. By pointing to it, the subject may report the comparison stimulus that provides additional (joint) control over the rehearsed hand sign. The pointing response (or a pressing response while maintaining the hand sign as shown in Figure 1) serves the function of an autoclitic by reporting to the listener something about the sources of stimulus control over the speaker's behavior.

The issue of sameness. Given this definition, let us recognize explicitly that to say a subject selects the stimulus that evokes the same topography as he or she is currently rehearsing is misleading. Such a usage encourages the invention of a hypothetical concept of *sameness* between topographies that ultimately carries the entire explanatory load (e.g., the sameness of stimuli selected is replaced by the sameness of the topographies the stimuli evoke).

In fact, the notion of joint control requires no judgment of sameness. Joint control occurs when a topography, already under duplic control, is evoked by some other stimulus. The subject is then able to report which stimulus allowed this to happen. And so, in the previous example, it would only be after finding that all six digits, as they were read from the page, allowed simultaneous duplic and tact performances that the subject would report having found the correct entry. Likewise, given the sample *mares eat oats* . . . and the various comparisons (e.g., *maizy oats* . . .), it is only by reading these comparisons while repeating the sample that one may be located that allows the duplic *mares eat oats* . . . to be repeated under joint duplic-textual control. Thus it is that the logical function of sameness is carried in this account by the behavioral function of joint control. Recognizing this, the word *same* may be used henceforth in order to avoid wordy locutions.

The origin and role of rehearsal. Although an emission of the sample topography is necessary for that topography to enter into joint control, its continuous rehearsal prior to that point is not. Thus, in locating one's own telephone number from a list of such numbers, it is unlikely that one would rehearse the number between attempts to match it with entries on the list. On the other hand, if the task involved a less well-known telephone number, the constant exposure to similar sequences might cause the reader to forget the sought-after number or confuse it with the sequences being perused. In this case rehearsals between attempts to match, by serving to maintain the requisite topography, would be differentially reinforced by a heightened selection accuracy. (Other contributions of the environment to the development and maintenance of rehearsal are discussed in the next section.) In general, and as with any other operant, the rate of emission of the rehearsed topography would be a product of the relevant contingencies of reinforcement. Contrary to Horne and Lowe (1997, p. 291), the notion of joint control does not require that there be constant rehearsals between the presentation of the sample and the ultimate stimulus selection: It only requires that there be one emission as each comparison is tested for its contribution to joint control.[4]

The Origins of Joint Control

A full account of the role of joint control in complex behavior requires a plausible description of its origin in the natural course of children's language development. The account proposed here does not dispute the commonly accepted course of language development, but only elaborates on the nature of the stimulus control that language exerts over other behavior (see Lowenkron, 1997, for a more complete discussion of this issue).

It is regularly observed that children learn three repertoires by the age of 2 years (Benedict, 1979). They first learn a receptive repertoire in which object names spoken by a caregiver serve to control selections, or other responses, to objects. This repertoire is taken as an indication of verbal comprehension. Later, after the minimal vocal repertoire has been acquired (up to several months later), children acquire two expressive repertoires: They *tact* the objects they have heretofore selected as well as the actions they have imitated, and they *echo* the words they hear.

What happens next is a matter of debate. Somehow, once the tact and echoic repertoires have been acquired, further practice with the original selection repertoire causes the three repertoires to interact so as to produce the complex, higher order, linguistic performances that characterize human behavior and that are the subject matter of this paper. Horne and Lowe (1996) have suggested that the three repertoires interact across the set of words a child has learned so as to produce the *naming relation:* an emergent, higher order relation that mediates the symbolic functions of language. Although several criticisms have been leveled against this account (e.g., Whitehurst, 1996), for present purposes the most relevant is directed at their assumption that novel names, and the objects they name, become related through the naming relation solely as a result of the covert rehearsal of names in the presence of the objects named (Horne & Lowe, 1996, p. 200). It may be argued, however, that just how this nondifferentially reinforced rehearsal serves to produce appropriate selection responding has not been fully explained (Horne & Lowe, 1996, p. 318; Lowenkron, 1996, p. 253).

On the other hand, how this sort of responding might be acquired under joint control is fairly straightforward. Thus, as Michael (1996) has suggested, as the child's environment becomes increasingly complex, selection accuracy in the initially acquired receptive repertoire is enhanced if children self-echoically rehearse the sample names they hear so as to preserve them while seeking the named object. If, while so rehearsing, the child encounters the named object, and if the child has previously acquired the currently rehearsed name as a tact for that object, then at that moment any further rehearsals of the name in the presence of the object would be under joint self-echoic/tact control. Were the child then to select the object as a result of learning acquired in the original receptive selection repertoire, a selection in response to the onset of joint control would be reinforced incidentally. Over the course of several such instances, given the generic nature of the joint control event described earlier, generalized control of the selection response by this event would subsequently emerge. As it became increasingly widespread in the child's repertoire, it could function to produce the many abstract performances to which we now turn.

The Logical Functions of Joint Control in Relations Between Words and Objects

The importance of joint control for a behavioral account lies in a single fact: It is fully defined by operant principles, yet it serves many functions typically ascribed to a variety of nonoperant, and frequently linguistic or cognitive, mechanisms. This section provides a behavioral account of some relations between words and objects that are generally described as symbolic or semantic: It is an account of word meaning phrased solely in terms of the interaction of simple verbal operants.

[4] Conceivably, when the topography is highly familiar (e.g., one's own name), there may be no rehearsal even at selection. However, the complex history that attends the selection of familiar stimuli makes selection under joint control hard to differentiate from unmediated selection, so the issue is probably best left to further experimental analysis.

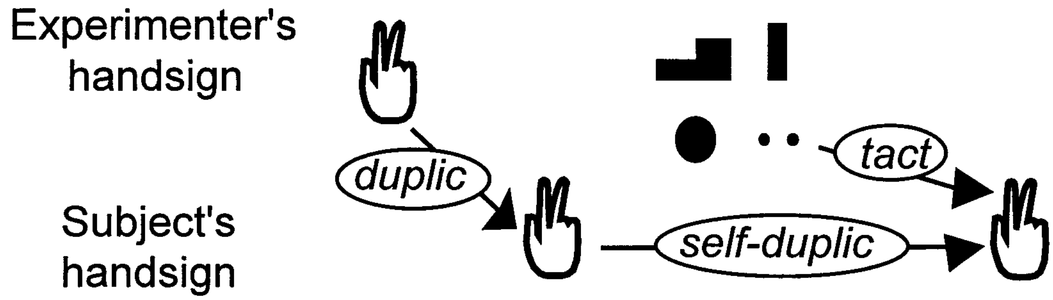

Fig. 4. An analogue to the specification of an object by a word. By modeling the hand sign shown, the experimenter has specified to the subject the two-dot comparison.

Word-Object Specification

As described above, where stimulus selection is unmediated, one stimulus is related to another only in the sense that the one causes a heightening in the probability of a selection response in the presence of the other as a result of their historical correlation with reinforcement. There is no disputing that this happens. But we do many other things with stimuli besides select them (e.g., report their existence), and often we do so without this sort of history. Thus, being told the name for something, or seeing a list of characteristics that describe something, may provide a history sufficient for one to say that the terms now *specify* the objects.[5] But what does it mean to say that a word specifies an object or event, or that a group of words *describes* something? Likewise, how do sentences specify logical and other propositions? And what is it that humans do that animals do not, that allows for all of this only in the former?

Given these questions, it is often argued that word-object relations involve emergent (e.g., symbolic or symmetric) functions outside the domain of simple verbal operants (Dugdale & Lowe, 1990; Hayes & Hayes, 1992; Horne & Lowe, 1996; Sidman, 1990). It is the premise of this proposal, however, that a complete account of the word-object relation, phrased solely in terms of the mechanical interaction of verbal operants, is possible.

Stimulus specification by the speaker's specification of a topography. The role of joint control in stimulus specification may be seen in the identity-matching example described earlier, but it is more clearly described in a hypothetical simplification of that experiment. Instead of presenting sample shapes, suppose that the experimenter had *told* the subjects which comparison to select on each trial by making the appropriate hand sign himself (Figure 4). And so, instead of the children making hand signs to the sample shapes as tacts, now their hand signs would be duplics: copies of the sample hand signs made by the experimenter. As a result, the task would devolve from identity matching to the selection of shapes in response to their names. Despite this, each sample hand sign can still be rehearsed by the subject as a self-duplic until a comparison is encountered that changes stimulus control over the hand sign from self-duplic to joint self-duplic/tact control.

In such a case, the experimenter's hand sign *specifies* the comparison stimulus in the sense that the experimenter's hand sign controls the topography of the subject's duplic hand sign, and the subject's hand sign, in turn, is allowed only by a particular comparison—the one comparison that initiates joint control over the hand sign. Thus the experimenter, by providing a topography for the subject to imitate, specifies the comparison stimulus the subject must choose.

The application of this analysis to vocal behavior is straightforward. In the task shown in Figure 5, Panel A, upon hearing the spoken phrase */black square/*, the subject, while rehearsing that phrase as a self-echoic, must then locate that object in the array. As with the hand sign example, the topography provided as the sample indirectly specifies the stimulus to be selected, namely, that stimulus

[5] Even without this minimal training, names may specify objects. Thus, the particular square specified by the phrase *largest square* depends on what is available and may change as squares are added or removed.

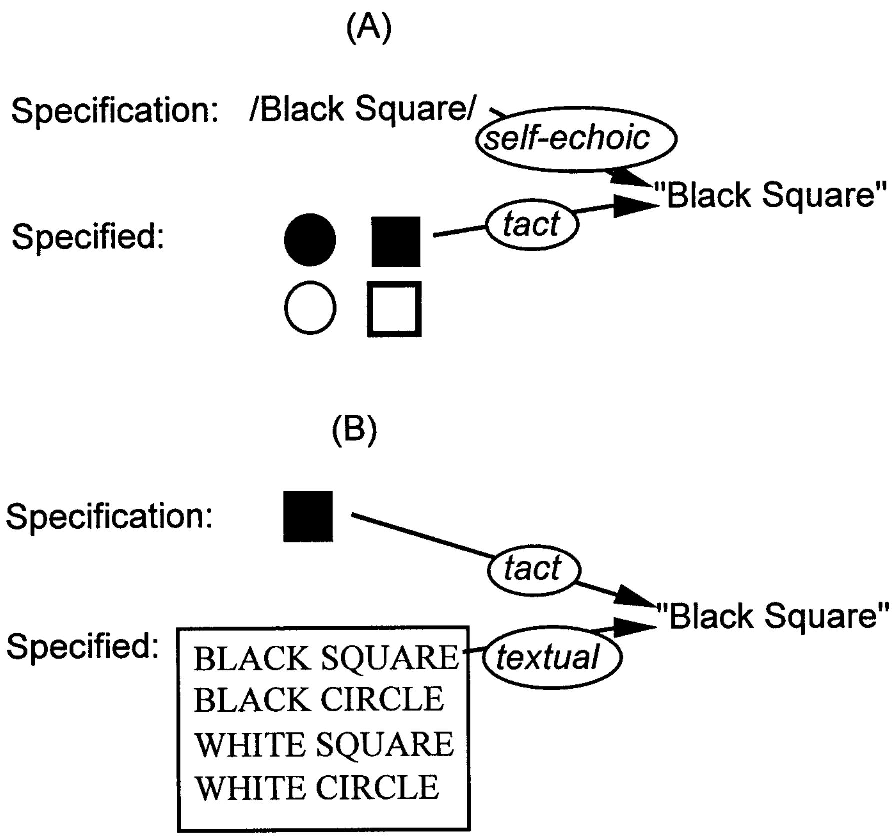

Fig. 5. Symmetrical word-object and object-word specification. Words heard by the subject are denoted by slashes; words spoken by the subject are in quotation marks; printed words are enclosed. (A) The spoken phrase "black square" evokes in the listener a self-echoic of the same topography. This topography can enter into joint control only with the comparison (the black square) that evokes the same topography (here, as a tact). Thus, the spoken phrase specifies a particular object by specifying the topography to repeat. (B) The black square evokes a tact that enters into joint control with the reader's textual production of that topography, thereby specifying the printed phrase that evoked the topography.

that evokes the same topography, as a tact, that the subject is currently rehearsing as a self-echoic. Thus the spoken phrase */black square/* does not directly specify (is not associatively related to) the black square itself; rather, the topography of that phrase comes under joint control only when it is evoked as a tact of the black square. It is this event that then evokes the selection response to the object.

In general, then, under joint control, one stimulus specifies another by means of the response topography it evokes: The topography evoked by the first stimulus determines the topography the second stimulus must evoke, and through this, specifies the second stimulus itself. A word, as a response topography, thus specifies a referent object by specifying the topography the referent must evoke. Put simply, a word specifies a referent because the referent itself evokes that word.

Stimulus selection by the listener's production of a topography. Complementary to the speaker's role in specifying a comparison stimulus by specifying a topography is the listener's role in supplying topographies in response to the comparison stimuli. Thus, as in Figure 5 (Panel A), prior to selecting, it is the listener who must tact the comparisons until the topography "black square" is evoked. Clearly, the listener is not a passive participant from whom selection responses are evoked, but rather is an active producer of behavior of which selection is but a part. In what follows, the pervasive role of the listener as a producer of behavior other than selecting stimuli will become increasingly evident and with it, the role of behavior analysis in the study of putatively cognitive processes.

Meanwhile, these considerations seem to suggest a solution to a broadly recognized problem, a problem behaviorists describe as the problem of word meaning for the listener and cognitive psychologists describe, to take a prominent example, as the *symbol-grounding problem*: the problem of how the response topographies known as words are ultimately related to the things they specify (Harnad, 1990, 1996). In the joint control account, the listener understands and agrees with the words specifying a referent object or event in the sense that they are also his or her specifications and topographies for that referent. The connection between a word and its referent is not an empty association, and given the word, the listener does not respond to the corresponding referent based on its physical features. Rather, the listener responds to the onset of the joint control that the referent initiates over the currently rehearsed topography of that word. A word is thus tied to a referent because they both can evoke a common topography.

But this is still not the whole story, because as we see in Figure 5 (Panel B), specification is bidirectional: A referent may specify a word in the same way that a word specifies a referent—by evoking a common topography. We now turn to a consideration of this and other aspects of the word-object relation.

Word-Object Bidirectionality

Ordinarily, it may be observed that along with the acquisition of names for objects and events, accurate selections of those stimuli in response to these names often appear spontaneously. Thus, in the prior example, immediately after learning to make hand signs to the transfer-set shapes (but not before), the children accurately selected those shapes (Lowenkron, 1988). This name-object bidirectionality (i.e., the symbolic function) is commonly understood to be a pivotal process in the acquisition and use of language, because it permits the efficient acquisition of much behavior without direct training (Horne & Lowe, 1996). It has thus been the object of considerable interest.

Early behavioral accounts of bidirectionality, based as they were on stimulus–response (S-R) relations, appealed to *backward associations* that were acquired while forward associations were being trained, as, for example, during paired-associate learning (Asch & Ebenholz, 1962; Ekstrand, 1966). More recently, a selection-based form of bidirectionality—symmetry—has been attributed to the innate sources that underlie stimulus equivalence (Sidman, 1990). Most recently, Horne and Lowe (1996) have suggested that unobserved rehearsal is the mechanism through which bidirectional responding emerges (but see Lowenkron, 1996).

Joint control and the emergence of bidirectionality. The current account brings with it a set of predictions as to the conditions under which bidirectional responding will emerge. Thus, when listeners have not yet learned to select under joint control but are selecting under unmediated S^D control, the acquisition of novel stimulus names should not produce bidirectional responding. Under these conditions, acquiring the tact relation illustrated in Figure 6 (Panel A) would not contribute to the selection of the stimulus in response to its name.

Things are dramatically different, however, when the listener has learned to select under joint control (Panel B). Now, after having learned to emit the spoken word "circle" in response to that shape (a tact), the word, spoken by another, specifies the selection of only that shape. More generally, after a listener has learned to select under joint control, *any* novel stimulus for which a tact or other stable topography is subsequently acquired may then be selected when that topography is evoked in the listener by the speech of another. Illustrating this, the precipitous emergence of generalized selection to novel stimuli after stable tacts to these stimuli were acquired has been reported across a wide variety of subjects and behaviors (Lowenkron, 1984, 1988, 1989; Lowenkron & Colvin, 1995).

A similar situation prevails with printed words. Saying a word in response to its printed version (a textual) provides a topography sufficient to allow the printed word to be selected in response to both its heard pronunciation and, if the word also functions as a tact, to the object it names (Lowenkron, 1991). Thus, assuming the subject can already read the word *circle* (Figure 6, Panel C),

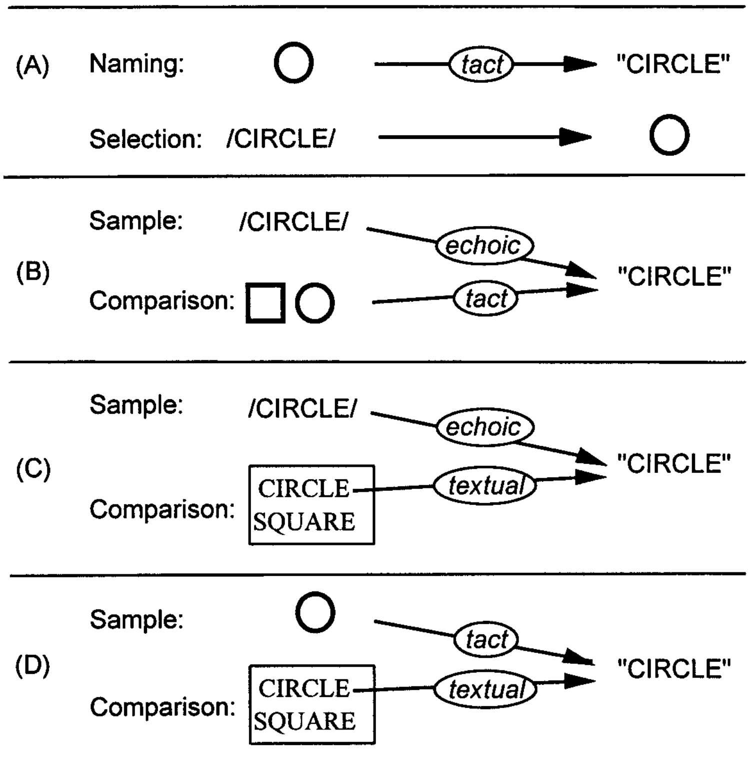

Fig. 6. The emergence of bidirectionality. (A) A naming response to a sample shape, and a selection response to a name. (B) Selection under joint control. The listener must select a shape in response to its spoken name. Here, the word *circle* enters into joint control with that shape. (C) Selection of the printed word in response to the spoken word. (D) Selection of the printed word in response to the object.

the listener's echoic pronunciation of the heard word *circle* provides a topography that comes under the joint control of one of the printed comparison words as it is read from the display. Similarly, in Panel D, the topography of a tact of the *circle* sample comes under the joint control of one of the printed-word comparisons when the word is emitted as a textual. In both of these cases, under joint control the printed word–spoken word (echoic or tact) relation is bidirectional (Lowenkron, 1991).

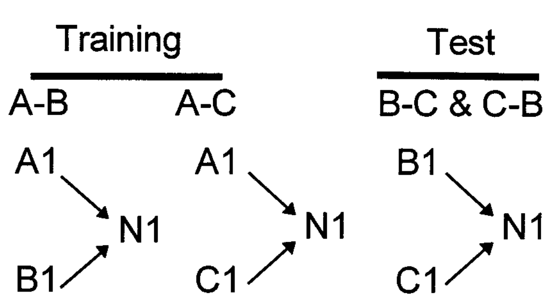

Fig. 7. The emergence of stimulus equivalence. Subjects creating their own names (i.e., N1) for stimuli in the course of A-B and A-C training is the sufficient condition for the emergence of stimulus equivalence under joint control.

Joint control and stimulus equivalence. If the training depicted in Panels B and C were to result in the emergence of the bidirectional responding illustrated in Panel D, such responding would qualify as stimulus equivalence (Sidman, 1990; see also McIntire, Cleary, & Thompson, 1989; Saunders, 1989). But under joint control, stimulus equivalence may also arise without explicitly trained names. Thus, if subjects devise and use their own names (N1) in the course of A-B and A-C learning (Figure 7) as some have claimed, then under the principles described above, these names would serve to mediate the relations defining stimulus equivalence (B-C and C-B) via joint control (Dugdale & Lowe, 1990; Eikeseth & Smith, 1992; Horne & Lowe, 1996). The relations that characterize stimulus equivalence are thus included within the larger set of relations described by joint control.

The Recognition of Objects and the Comprehension of Descriptions

A pigeon selecting red squares or green diamonds under the control of the spoken sounds ''carré rouge'' and ''losange verte'' would no more be said to *comprehend* the meaning of those words than would a person with no background in spoken French engaged in the same task. For both, the task would only involve unmediated selection, and the spoken sounds would be no more than conditional stimuli. But there are important differences here. Ultimately, the person, but not the pigeon, may be trained to comprehend the words so that appropriate responding would appear even with novel phrases naming stimulus properties and describing complex relations (e.g., *red diamonds, but not green*). These observations raise many questions: What does it mean to call a stimulus a description? How does a description, heard or read, differ from an S^D? How are the things described recognized from their description? Why is a person, but not an animal, able to comprehend a description? And indeed, what does it mean to talk of comprehension?

In some cognitive accounts, the listener serves solely as a recipient and processor of the information contained in speech (e.g., Jackson & Morton, 1984). A description is considered to be a packet of information, and its comprehension is left to hypothetical internal processes of perception, parsing, and information utilization (Anderson, 1990, p. 360). This approach does not solve the problem, however, because ultimately, at some level, the subject must respond to (i.e., interpret) the information. Moving this process inward only makes it less accessible.

As to the behavioral account, there are several possibilities. According to one account, if the subjects were to learn to match both the objects and their printed descriptions to spoken phrases, and then spontaneously match the objects and printed words to each other, this would show the emergence of both auditory comprehension and reading comprehension (Sidman, 1990, pp. 94–96). But even if such responses were also trained to the spoken phrases ''losange rouge'' and ''carré verte,'' it would not necessary follow that the subjects would then emit the tacts *rouge* and *verte* to the appropriate colors and *losange* and *carré* to the appropriate shapes (Goldstein, Angelo, & Whetherby, 1987). And so, these phrases would not be describing the stimulus features in the way they do for a native speaker, nor would subjects be comprehending these phrases in the way a native speaker does. As illustrated below, the inadequacies of this account for addressing comprehension are further compounded when the material to be comprehended contains prepositional, syntactic, or logical relations (e.g., green *or* square).

The production principle. In another account, Skinner (1957, p. 268) comes close to the issues raised here in his discussion of supplementary stimulation. He points out that when the controlling variables in a situation have a common effect on the listener and speaker, verbal stimuli emitted by a speaker may evoke similar behavior in the listener already at a high level of strength (although not so high as to be emitted), such that the listener appears to be following or comprehending what is said. As a result, Skinner points out, the listener ''completes a sentence for the speaker if his own behavior is more rapid, or if the speaker is for any reason interrupted. He joins with the speaker in emitting an important word or phrase. Even when he does not emit the response, he may recognize his own participation by saying 'He took the words right out of my mouth' '' (p. 269).[6] Thus, Skinner suggests, comprehension requires a common vocabulary with respect to the situation being discussed, or to put it in the current terminology, comprehension requires evocation in the listener of common topographies by the speaker and jointly by the situation being discussed.

Enlarging upon Skinner's analysis, the joint control account gives primacy to the comprehending listener as an active behaver rather than as a receiver. Viewed through this lens, the comprehension of words, syntactic and

[6] This response seems to be an autoclitic that reports the change in response strength upon hearing the words of another.

logical relations between words, and the terms describing such relations all depend upon the ability and readiness of the listener to emit these same words and phrases under appropriate stimulus control, and, optionally, to report this joint control autoclitically as understanding. In principle, the receptive function of language comprehension is premised on the expressive: In order to behave in the manner generally ascribed to comprehension, the listener must first respond appropriately to the stimulus as a speaker. Comprehension is based on production.[7] As we shall see, this *production principle* has the salutary effect of behavioralizing the entire notion of comprehension by eliminating the metaphor of information reception, thereby moving the processes of comprehension from the psychic interior to the behavioral surface.

The primacy of the expressive function over the receptive has been argued by Lowenkron (1991) with regard to a distinction proposed by Michael (1985) between *selection-based* verbal behavior, where subjects point to or otherwise select stimuli in response to verbal stimuli, and the ordinary spoken and written *topography-based* verbal behavior described by Skinner (1957). Essentially, the argument echoes those made above. Thus, in the selection-based intraverbal, where a printed word or phrase must be selected in response to a spoken word without a history of prior reinforcements for such selections (e.g., *Deng Xiaoping*), accurate selection depends on the listener emitting the appropriate topographies to the printed phrases by reading them (a textual) and then selecting one that permits an emission of the topography under joint self-echoic/textual control.

Likewise, in manded stimulus selection, where an object must be selected in response to a verbal stimulus, a selection response to a novel object in response to its name (e.g., *black square*) does not arise in direct response to that verbal stimulus. Rather, when stimulus selection is under joint control, the listener selects the object that evokes, as a tact, the topography also emitted as a self-echoic repetition of the mand. Again, selection depends on a prior production.

Recognition from a description. Experimental support for this interpretation with vocal responses was recently demonstrated with 6- to 7-year-old children (Figure 8). In the first phase of the experiment, the children learned to select six stimuli in response to three-word samples that named the colors, shapes, and border patterns of the stimuli in unfamiliar terms (e.g., *king-flag-clip*). The acquisition of this behavior was very slow, and once learned, the children could not tact the six stimuli by saying the complete three-word names despite the many trials in which they had heard these names and had selected the stimuli in response to them. Nor, in a subsequent test for generalization, could the children select novel stimuli, composed of new combinations of the same colors, shapes, and borders, in response to new combinations of the names. The slow rate of initial acquisition, along with their inability to either tact the stimuli or select novel combinations of the stimuli in response to novel combinations of names, suggests that comparison selection, up to this point, was simply unmediated stimulus selection: Comparisons were selected only on the basis of their relative probabilities of evoking the selection response.

In the next phase, the children were trained to tact each of the values of the color, shape, and border dimensions as each value was shown separately, but this did not improve performance in the generalization test. However, after the children learned to tact these values as they existed within the actual stimuli, accurate generalized matching appeared immediately. Thus, accurate selection of the stimuli depended on the accurate production of tacts to the stimuli. It is not unreasonable to attribute this sudden improvement in accuracy to the fact that it was only at this point that the children could select comparisons under joint control. Now, given a three-word sample, they could rehearse it as a self-echoic until they encountered a comparison stimulus that allowed a further repetition of the sample topography as a tact so that the topography was emitted under joint self-echoic/tact control. Thus, accurate generalized selection depended upon the production of accurate tacts.

The foregoing suggests that the spoken

[7] This notion closely parallels the motor theory of speech perception: How speech sounds are heard depends on how they are produced (Liberman, Harris, Hoffman, & Griffith, 1954).

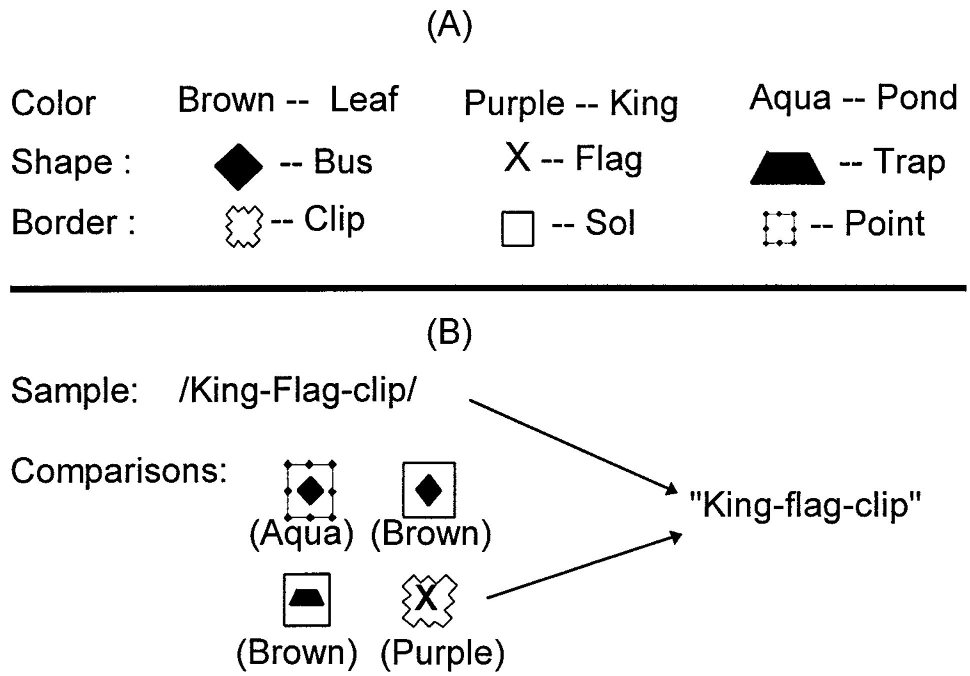

Fig. 8. Generalized selection from a description. (A) The colors, shapes, and borders used to construct stimuli and their names. (B) Typical stimuli constructed from colors, shapes, and borders. The colors of the stimuli are indicated in parentheses. Here, the stimulus comprised of the elements named *king, flag,* and *clip* is selected under joint control in response to its description.

samples played two different roles. In the first phase they only evoked unmediated selection responses. Later, they contributed, along with the subject-produced tacts, to joint control, which in turn evoked stimulus selection. In this case the spoken samples appeared to serve as *descriptions* of the stimuli: They were topographies that did more than just evoke a selection response to the stimuli, the same topographies also served as the listener's own tacts for the stimuli.

Given this, we may distinguish between a conditional stimulus and a description. A conditional stimulus is a stimulus, possibly verbal, in whose presence a selection response to a comparison becomes more probable solely as the result of a history of reinforcement for such selections. This control is not dependent on the listener's ability to reproduce the stimulus, and conditional stimuli foster only limited generalized responding through primary stimulus generalization. A description, on the other hand, is always a verbal stimulus. It specifies some other stimulus object or event by contributing joint control to the subject's own tact of that other stimulus. Thus, in order to be able to respond appropriately to a description, especially a novel description, one must be able to emit the same topography as a tact. When verbal topographies acquired as echoics come to function as tacts, their roles change from conditional stimuli to descriptions.

By this definition what may be a description for one subject may not be for another, depending on what the subject already knows in the sense of what the subject can tact. Thus, if listeners who had learned to respond to the words *rouge, verte, carré,* and *losange* as conditional stimuli were asked what they were doing, they might describe the contingencies, but they would not say that the words described colors and shapes. The words would change from conditional stimuli to descriptions, however, after the subject learned to emit these words as tacts; now, the topographies evoked as tacts by the colors and shapes could come under joint control with the subjects' own self-echoic repetitions of these topographies.

This distinction between a description and a conditional stimulus also clarifies the com-

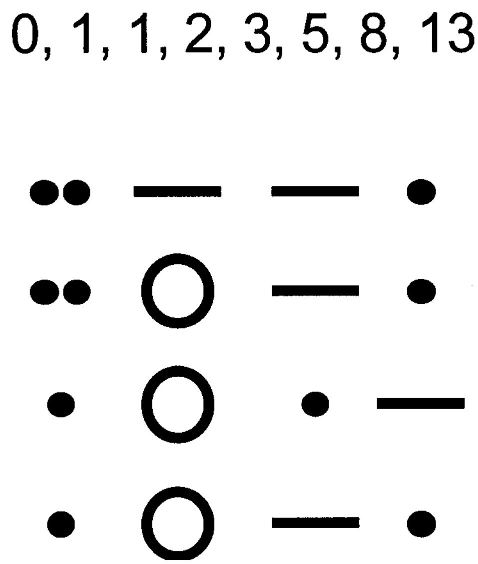

Fig. 9. Comprehension of a description and recognition of a described stimulus. The number series is a Fibonacci series, but this description is comprehended only if one has that phrase available as a tact. One series of shapes is recognized from the description *point-oh-bar-point* only when those terms enter into joint control with the tacts corresponding to the elements of the series.

monplace notions of *comprehension* and *recognition.* Thus, to say someone *comprehends* a description is to say that the topographies evoked as self-echoics by a description may also serve as tacts for that person. Consider the numbers shown in Figure 9. It is a Fibonacci series. One comprehends the phrase *Fibonacci series* as a description of these numbers only if that phrase is available as a tact for a number series constructed of adjacent sums. As for the term *recognition,* on this account the event we call *recognition from a description* may be identified with the actual moment of onset of joint control. Thus in Figure 9, the stimulus described as *point-oh-bar-point* is recognized at the moment (and only at the moment) that the rehearsed topography occurs under the joint control of that stimulus. Subsequent references to its location would not involve selection under joint control unless its location were forgotten, and the subject once again had to look for and find *point-oh-bar-point.* (Clearly, further study is needed to learn how we apply these unfamiliar usages to stimuli, and how such applications of approximate descriptions may affect our perception of the stimuli.)

The comprehension of a description, and the recognition of a stimulus from its description, are thus complementary. Comprehension permits recognition. For a description to be comprehended by a particular subject, its terms must be capable of functioning as tacts for that subject. And when they do so with respect to a particular stimulus, that stimulus is recognized as the one described.

Confirmation: The case of one comparison. Closely related here are situations in which a selection response is not required because there is but a single comparison stimulus. In these cases concern focuses on the nature of the relation between the description and the stimulus described, and the autoclitic responses to the joint control event are the topographies *yes* or *no* rather than a selection response. Thus the query *Is it a blue square?* evokes an autoclitic *yes* or *no* depending on whether or not the object described evokes a tact that can enter into joint control with the self-echoic of the query.

Something very similar seems to occur when we check ourselves. Usually, in this case, a tact of a sought-after condition is rehearsed as a self-echoic and checked for agreement against a given stimulus. Thus, we may re-add a list of numbers to see if there is an identity match with the first sum. (The interaction of tact and self-echoic becomes more apparent if the two sums are not on the same page.) Though not pursued here, it may well be that behavior of this sort plays a role as large (or larger) than selection-based behavior; consider how many times we respond "yes" or "no" throughout the day. Certainly, not all of this responding is under joint control, but a significant proportion may be.

Prepositional, syntactic, and logical relations. Descriptions generally consist of more than nouns and adjectives, but the principles of comprehension described here seem equally applicable. Consider the role of prepositions and syntax. The phrases *blue on red* and *red on blue* describe different things due both to their respective word orders (i.e., syntax) and to the nature of the preposition they both contain. But the comprehension of these phrases as descriptions, in the sense described here, depends only on the subject's capacity to produce the topographies appropriately. And so, when faced with a stimulus containing red above blue, its selection in response to the description *red on blue* would depend on the subject tacting the stimulus with a response of the same topography. Here, accurate selection depends on a prior history of appropriately reinforced tact productions with respect to the preposition *on*.

Descriptions may also contain elements specifying logical relations as in *red or blue,* or *red and blue,* or *if red, then blue*. These logical elements serve as instructions concerning the proportion of instances of joint control that must exist. Thus the term *or* specifies that an object that allows joint control with either color mentioned in the description is sufficient to evoke its selection, whereas the term *and* indicates that the color terms on both sides of the word *and* must enter into joint control. Just how logical terms operate in this manner remains to be explored, but the larger implication of this argument is that the comprehension of terms describing relations between stimuli depends on one's ability to respond to actual instances of the described relation with appropriate tacts.[8] Again, comprehension presupposes production.

Finally, we may observe that descriptions are often interpreted, and differences in these interpretations affect responding. There are at least two possibilities here; interpretation may involve changes in the self-duplic leg or in the tact leg of joint control. The first case more generally typifies what is meant by interpretation. Thus, the phrase *the good ones are blue and white* might lead to the selection of items blue *or* white, or to the selection of items blue *and* white depending on whether the self-echoic is rehearsed as a single phrase *blue and white* or as two phrases, *blue* and also *white*. A case in which interpretation arises in the tact leg may be seen when a person regularly tacts rectangles as *squares,* because then, in response to the sample *square* he or she may select both rectangles and squares.

Interpretations, of course, are not immutable. Changes in the name given a class of stimuli will affect responding. Sometimes this may happen directly. Thus, learning to tact squares and rectangles appropriately would affect subsequent selection behavior. But there are many less obvious examples of how learning in one situation may affect tacting and responding in another. Thus, learning to emit the tact *fruit* in response to anything with seeds in it makes it more likely that a person will select a tomato in response to the sample *fruit* even if fruit and tomato were never explicitly paired. There are a good many variations here that might be tried experimentally.

We may summarize all of the preceding by simply noting the need to appreciate the role played by subjects' responses prior to comparison selection, because when subjects are responding under joint control, they will act appropriately to phrases describing complex logical and prepositional relations only to the extent that they are themselves able to produce appropriate topographies. And in doing so, they may be said to have responded to the phrases as descriptions, comprehended and interpreted these descriptions, and then recognized what was described.

The Logical Functions of Joint Control in Relations Between Objects

In contrast to word-object relations, where the characteristics of the words and the objects they specify bear no formal relation to each other, there exists a large class of performances based on consistent relations between the formal properties of objects. The capability to respond to these relations not only frees behavior from the particulars of the situation in which it was acquired but also broadens responding beyond the narrow limits allowed by primary stimulus generalization. Relational responding allows responding to generalize across situations that have no physical aspects in common except for the relation between their elements. Thus, selecting stimuli *larger than, smaller than, before,* or *after* other stimuli or selecting stimuli identical to, or nonidentical to, other stimuli all involve responding on the basis of relations between the primary stimuli rather than on the basis of any characteristics possessed by the primary stimuli themselves. This is especially true when the stimuli are not present simultaneously. Typically, relational responding has been ex-

---

[8] With complex, multipart stimuli, accounts of unmediated stimulus selection appeal solely to the accumulation of stimulus parts to evoke a selection response. Thus terms like *red* and *blue* act cumulatively to evoke a selection response to a red and blue stimulus. But such an account cannot explain how the further addition of a relational term (e.g., *on*) specifies one comparison (red on blue) as opposed to another (e.g., blue on red).

plained in terms of concepts or rules (Bucher, 1975; Zentall, Edwards, Moore, & Hogan, 1981). More recently, Hayes (1991) has suggested that such responding might be described in terms of the relations entailed by training.

Responding to Relations Through Joint Control

As discussed earlier, the joint control event does not depend on any characteristics peculiar to the primary stimuli. Rather, the stimulus characteristics of the joint control event are general to any case in which a second stimulus evokes the topography currently being rehearsed in response to the first stimulus. As a result (at the risk of a slight oversimplification), the characteristics of all possible stimulus pairs fall into two categories: those stimulus pairs that do and those pairs that do not enter into joint control. Pairs of the first type may then evoke one autoclitic response (e.g., pointing, pressing, or saying), and pairs of the second type may evoke a different autoclitic or none at all. When there is only one autoclitic, the subject may respond to instances of a relation and ignore noninstances, or vice versa. When there are two different autoclitics, the subject may classify stimulus pairs by responding differentially to instances and noninstances of a relation.

As a result of this process of stripping away the characteristics peculiar to the primary stimuli and translating all instances and noninstances of the currently relevant relation into instances and noninstances of joint control, nothing is left to impede generalized responding to these relations when they occur with novel stimuli so long as these stimuli evoke accurate and stable (unchanging) response topographies that may then enter into joint control. As we shall see next, this stripping process provides an exceedingly powerful basis for generalized relational responding across a wide range of relations. When it is combined with precurrent behavior that transforms stimuli before they enter into joint control, the result is an even wider variety of performances, some with characteristics that seem goal oriented.

Identity matching. The fundamental formal relation between stimuli—identity—naturally involves the simplest set of behavioral relations. Thus, as we have seen (Figure 1, Panel B), where identity matching develops under joint control, the sample and the matching comparison within each stimulus set act to produce a common hand sign which, at some point, occurs under joint control and evokes a selection response. Identity matching under joint control is thus a purely mechanical behavior process. The sample, by evoking a particular hand sign (or word), necessarily specifies the identical comparison: the one that allows the sample hand sign (or word) to remain unchanged as it comes under joint control. Identical stimuli (especially when presented successively) thus need not be picked because of their perceptual identity or their formal identity, but because they evoke a common hand sign or a common word or description. Nothing in this process distinguishes familiar stimuli from novel ones: Any stimulus that consistently evokes the same topography may be matched to itself solely on the basis of the joint control it engenders. As we see next, other relations are mechanical derivatives of this process.

Similarity. Objects selected in response to other objects may differ on some dimensions and thus be similar rather than identical to the sample. One way this may happen is if the sample object is not tacted fully or appropriately. Thus, as illustrated in Figure 10, a sample consisting of a two striped squares may evoke only the tact *square,* eventuating in the selection of a square, or it may evoke the tact *striped square,* producing a selection that disregards number but not shape and pattern.

In general, in cases of this sort, both instructions and context play a role. Consider a person's responses with respect to a set of dishes. In the kitchen, the term *identical* as in *get me a plate identical to this one* would serve to insure that all dimensions of the sample plate were tacted by the listener, and thus controlled selection of the comparison, whereas the instruction *get me one like this* might serve to loosen the sample's control over the initial tact, resulting in the listener's tact *big plate* and the selection of a plate that was big but was decorated with a different pattern than the sample. As for the role of context, consider what would happen if the request *get me one like this* were made while setting the table: It would more likely result in the selection of a plate identical to the sample rather than similar to it.

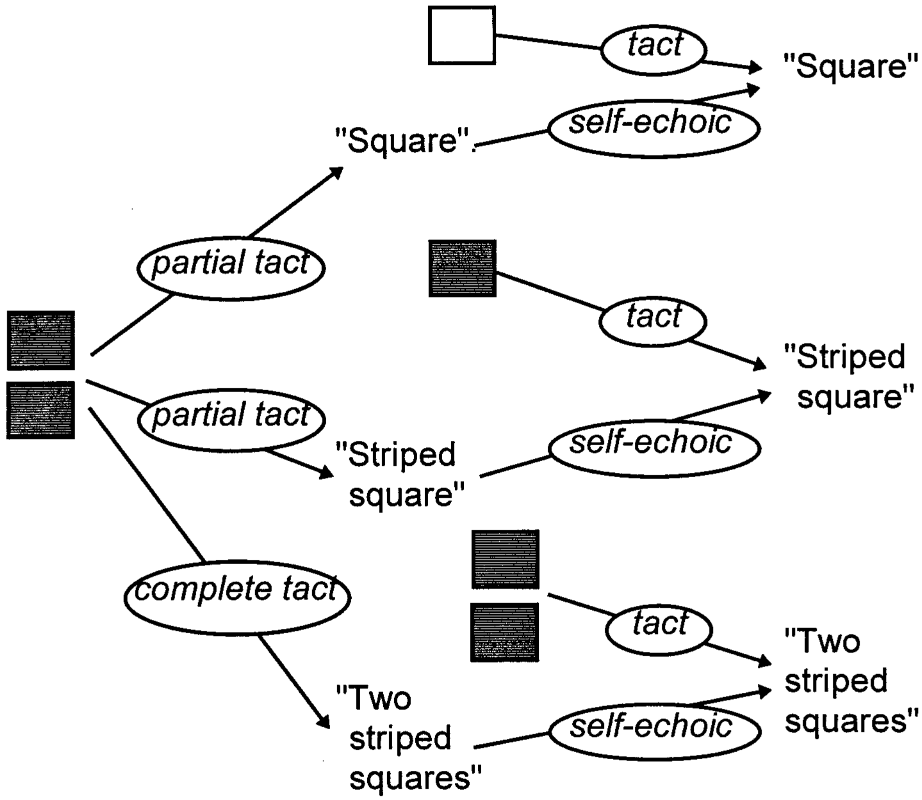

Fig. 10. Variation in similarity. The sample may be responded to with tacts of varying completeness leading to the selection of stimuli that are either similar or identical to the sample.

Selection based on similarity may also arise as a result of primary stimulus generalization interacting with joint control. As the result of such generalization, a large turquoise cup might evoke the tact *blue bowl,* thereby resulting in the selection of an item under joint control that was only similar to the sample.

Clearly, there are many unanswered questions here. Thus, just as the logical relations discussed earlier raise questions about the nature of stimulus control over the comparison tact, so here too do questions arise as to how instructions and context enter into the control of sample tacts. The interaction between joint control and primary stimulus generalization also remains to be explored.

Responding to Relations Through Transformational Responses

In contrast to the above, joint control may also occur based on features the stimuli do not hold in common. Thus, we are often asked to select stimuli that are *larger than, wider than, before, after,* or *further from* some sample. The nature of responding in these cases has been analyzed in several studies (Lowenkron, 1984, 1988, 1989; Lowenkron & Colvin, 1995). We shall examine this next.

Consistent relations. Unique to the kind of behavior discussed here is the role played by precurrent *transformational responses.* These responses are operants. They serve to transform the topography of the response emitted to the sample so that a different topography is rehearsed as a self-duplic, and thus a comparison different from the sample is specified for selection. Transformational responses allow the comparison selected to bear a consistent relation to the sample. For example, in alphabetizing a series of names, as we complete the search for names beginning with one letter, we intraverbally transform the rehearsed topography to the next letter and begin to search again. By repeatedly transforming, we alphabetize the names completely.

Although transformational responses are

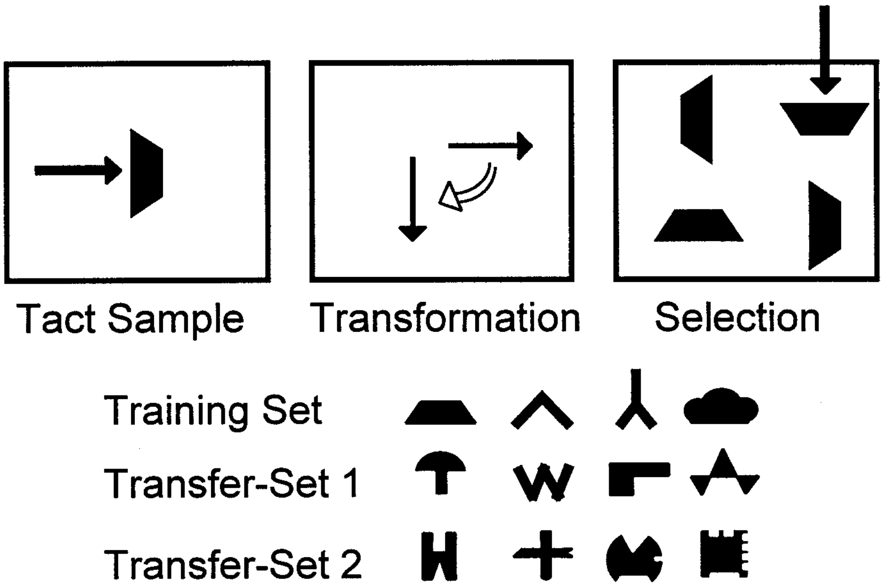

Fig. 11. Components of the orientation matching task. After setting the arrow to tact the orientation of the sample, the subject transforms the orientation by turning the arrow 90° clockwise and then selects a comparison appropriate to the new orientation. Transfer Set 1 contains three symmetrical and one asymmetrical shape, and Transfer Set 2 contains three asymmetrical and one symmetrical shape.

generally vocal, they need not be (Lowenkron, 1984, 1989). Thus, Lowenkron (1984) reported a study designed to measure transformational behavior directly: In this study nursery-school children learned to align a hand-held arrow with the axis of symmetry of each member of the set of training shapes shown in Figure 11. They then practiced this performance when these shapes appeared as samples in a matching task. (These responses were nonvocal tacts of the shapes' orientations.) The children also learned to transform the sample tacts by turning the arrow 90° clockwise from its initial orientation as a tact, and learned to select the comparison that was in an orientation that could be tacted by the new orientation of the arrow. Later, when new symmetrical stimuli were presented (Transfer Set 1), all components of the performance generalized: The subjects continued to represent the orientations of the samples' axes of symmetry with the arrow, continued to transform these orientations, and continued to select comparisons appropriate to these transformations. Subsequently, however, when the asymmetrical shapes of Transfer Set 2 were presented, they were not tacted consistently and selection accuracy fell to random levels. Later, after the children were trained to tact the orientations of the new shapes, selection accuracy immediately rose to high levels. Both the transforming and the comparison-selection behavior generalized to these stimuli without additional training. Thus, neither the transformational responses nor (as discussed previously) the comparison selection responses were under the control of features specific to the stimuli; rather, they were under the control of features common to all trials.

This example shows how transformational responding may interact with comparison selection under joint control to produce generalized relational responding. The transformational response, controlled by general task features, serves to change the topography of the self-duplic from what is initially evoked as the sample tact. In changing this topography, it changes the comparison specified by the topography, thereby causing the selection of a comparison bearing a particular relation to the sample.

Instructing consistent relations. One more level of complexity must be considered. Like any other class of operants, transformational responses can be brought under stimulus con-

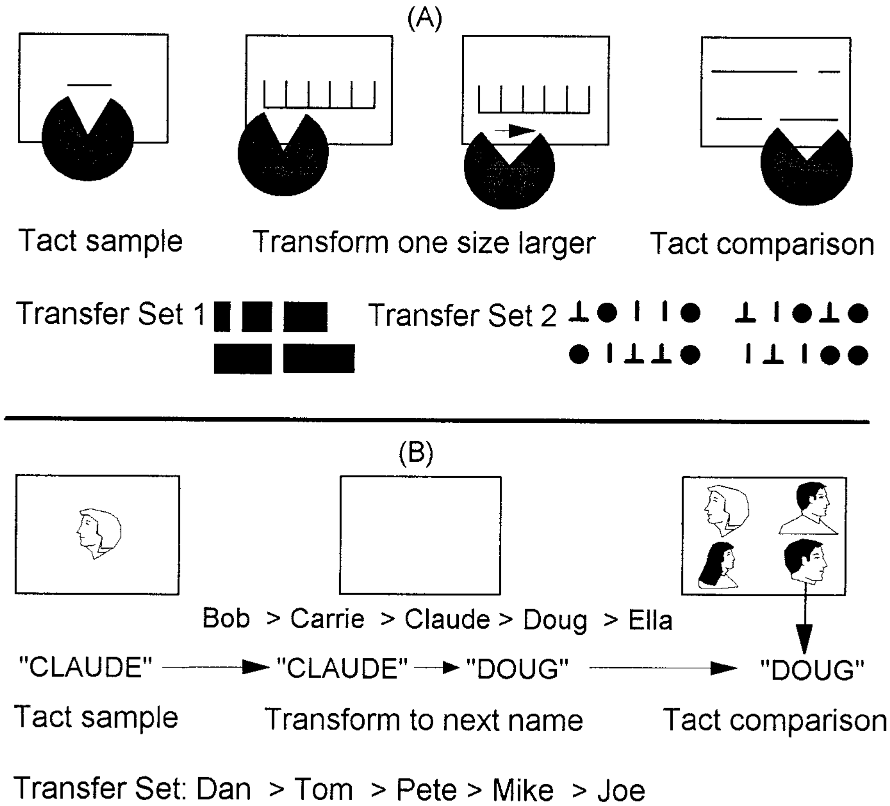

Fig. 12. Two matching tasks under instructional control. (A) The subject first adjusts the measuring device to fit the length of the sample line, then readjusts the device under the control of the current background color. Here, the subject has enlarged the device one size, and so will seek a comparison one size larger than the sample. Shown below are the rectangles and line-and-dot patterns used in tests for transfer. (B) The subject must name the sample, rehearse the name over the delay, increment the name rehearsed if the background color demands it, and then select the comparison appropriate to the new name. The names, but not the faces, of the transfer set are shown below.

trol. As a result, performances involving different relations may be produced by presenting stimuli that control different transformational responses. These stimuli are referred to here as instructional stimuli. Their role is illustrated in a study reported by Lowenkron (1989) and by data from other studies conducted in my laboratory.

Children (4 to 5 years old) were taught to select lines that were longer than, equal to, or smaller than a sample depending on whether the background color for that trial was green, blue, or red (Figure 12, Panel A). To produce this performance, the children were first trained to set a measuring device to fit the ends of the samples (horizontal line lengths) and then to change the setting of the device under the control of instructional stimuli: They set the device one size smaller if the background was red, left it unchanged on a blue background, or made it one size larger on a green background. The subjects then selected the comparison line length that fit the transformed setting.

The instructional stimuli maintained control of generalized matching when novel stimuli were substituted for the training set. Thus, when rectangles of varying widths were substituted for the horizontal line lengths (Figure 12, Panel A), generalized instructional control and generalized matching continued. However, when the dot and line patterns of Transfer Set 2 replaced the rectangles, selection accuracy fell to random levels until the subjects were trained to set the device to fit the distances between dots. And so, as with the asymmetric stimuli in the orientation-matching task, here too the absence of a tact specific to the relevant dimension prevented generalized responding. But again, once this behavior was acquired, all the other components generalized, permitting immediate generalization of the entire performance.

The instructional control of a vocal transformational response is illustrated by data, some of which were reported by Lowenkron and Colvin (1995), from a study in which children learned to name a series of faces (Bob, Carrie, Claude, Doug, Ella) in a constant order (i.e., as an intraverbal) (Figure 12, Panel B). They also learned to emit this intraverbal under the instructional control of a green background but not a blue background. As a result, on green and given a sample name, selections of the face whose name appeared next on the memorized list (i.e., *after* the sample name) were reinforced; on blue, selecting the named face was reinforced, and on red (with no additional intraverbal training) selecting the face one back from the sample name (i.e., *before*) was reinforced.

In a subsequent phase, a transfer set containing new faces was presented, and their names were taught. Subjects were then taught to recite these new names in a fixed order (Dan, Tom, Pete, Mike, Joe). As in the prior case, instructional control over the transformation, and thus over the matching relation, generalized to novel stimuli: Here, depending on the background color, subjects selected the face one forward, the same as, or one back from the name given. But there was something unique here. In the studies with arrow orientation and line length, the topography of the transformational response remained the same from the training set to the transfer sets: turning the arrow or changing the setting of the device. Now, however, with the transfer set the subjects incremented a new list of names as the colors dictated, despite the fact that they had no training in doing so with this list. The transformational responses of incrementing and decrementing an intraverbal thus generalized to these novel names with no specific training, and with it generalized stimulus control of selection by serial order!

As discussed below, generalized transformational responding takes on crucial importance in describing those patterns of behavior typically ascribed to goal orientation. However, both the nature and the extent of these contributions need further analysis. Also needing study is the question of how instructional stimuli actually serve to evoke a new topography. Thus, if a sample evokes the third name on the list, how does the instructional stimulus cause the subject to increment to the fourth name? What is this generalized transforming?

Even without knowing how instructional stimuli work, the foregoing studies still demonstrate how generalized instructional control over transformational responding may interact with joint control. An instructional stimulus does not actually specify a relation between stimuli; rather, it controls the transformation from one rehearsed topography to another. This topography in turn specifies which comparison is chosen and thus the nature of the relation between the stimuli. Instructional control of relational matching is thereby reduced to simple discriminative stimulus control over an operant. As we shall see next, although this is a parsimonious account of a highly abstract performance, it is not an oversimplification: It not only describes the conditions under which a subject may engage in this behavior but also describes the conditions under which a subject may report what he or she is doing.

Reporting Relations: The Role of the Autoclitic

Up to this point, the discussion has largely dealt with performances in which there was but a single autoclitic response to the onset of joint control, generally a selection response such as pointing. But the susceptibility of other verbal topographies to autoclitic control allows speakers to report about other aspects of joint control beside its onset. After discussing the nature of autoclitic responses, some implications of this extension are considered.

According to Skinner (1957), autoclitics, like other verbal operants, are defined by both their antecedents and their consequences. As to the antecedents, descriptive autoclitics are controlled by events and by relations controlling other features of the subject's verbal behavior (p. 313). These may include "private stimuli associated with vocal behavior, possibly of a covert or even incipient form'' (p. 314). Presumably, these private stimuli can include stimuli associated with the onset of joint control. But if the onset of the joint control event can evoke autoclitic responding, then so should the onset of other events as, for example, when a comparison is encountered that evokes a topography that conflicts with the currently rehearsed topography (Lowenkron & Colvin, 1992; Skinner, 1957, p. 322). By responding to both of these events (the onset of joint control over a common topography as well as the evocation of conflicting topographies), both the presence and the absence of the stimulus specified by a sample may be reported.

As to the defining consequences of the descriptive autoclitic, the function of this operant is to modify the effect (on the listener) of the verbal behavior it accompanies. When it is under joint control, a selection response indicates to the listener which comparison stimulus enters into joint control with any self-duplic currently being rehearsed by the speaker. Thus, as with any descriptive autoclitic, the selection response indicates to the listener a source of stimulus control over the behavior of the speaker that is not otherwise evident.[9]

Typically, of course, the autoclitic is a vocal topography. As the following demonstrates, through such autoclitic reports speakers can describe many complex and abstract features of the stimuli, and the relations between stimuli, that control their behavior.

Reporting word-object relations. Earlier, it was shown that joint control provides a mechanism by means of which words and objects may mutually specify each other. We now focus on the means by which a speaker may report this mutual specification. A speaker shown the printed word *book* and an actual book will report that the word is the name of the object and that the object is named by the word. Each of these is an autoclitic report occasioned by the onset of joint control between the textual evoked by the printed word and the tact evoked by the book itself. By saying that the printed word is the name of the object, the subject is reporting that the topography first evoked under textual control by the printed word comes under joint control of the tact evoked by the object itself. Conversely, in saying that the book is named by the word, the speaker is reporting that the tact evoked by the object enters into joint control with the self-echoic rehearsal of a textual. (For the sake of simplicity in this and subsequent examples, let it be understood that rehearsals of one or the other verbal behavior provide the self-echoic component, depending on which stimulus the subject looks at first—in the present case either the textual or the tact might be repeated as a self-echoic depending on whether the subject looks from the object to the word or vice versa.)

In a similar fashion, with a set of words (i.e., *small red chair*) and the corresponding object, the occurrence of joint control not only would allow the presence of one to evoke the selection of the other but also would provide the event that allowed the subject to report (i.e., tact as an autoclitic) that the words described the chair, or that it is the chair that is described by the words.

Reporting identity and similarity. When both of the stimuli are objects, there may be reports that the objects are *identical to* or *the same as* each other. Three or more objects that sequentially evoke the same tact may then evoke the autoclitic tact ''They are all the same.'' If there is only partial joint control along with conflict (e.g., a small red chair and a small blue chair), there may be a report that the objects are *similar.* There are, of course, many relations other than identity that may exist between stimuli, and as we see next, there are many ways in which they may be reported.

Reporting relations as intraverbals. For the sake of completeness, we begin with a simple case that does not involve joint control. It is illustrated in the two experiments described earlier in Figure 12 (Lowenkron, 1989; Lowenkron & Colvin, 1995). In the first of these, the children learned to enlarge or reduce a measuring device, and in the second they learned to increment a list of names forward or backward. In both cases these transformational responses were under the control of the screen colors. The colors thus served as instructional stimuli that controlled the nature of the overall matching relations. Al-

[9] Alternately, the speaker may not emit the autoclitic response, thereby producing the case in which the speaker recognizes the specified stimulus but does not so indicate to others. Unlike unmediated selection, under joint control recognition is separate from selection.

A) 2 - 4 - 6 - 8

B) E - D - C - B - A

C) R - T - U - V - S

Fig. 13. Three relations between the members of stimulus sets.

though it was not done in these experiments, these colors could have been brought to control other verbal responses as well, such that the green screen evoked the intraverbal responses *larger* or *next forward* and the red screen evoked the terms *smaller* or *next back* (Skinner, 1957, p. 99). In this case reporting that the selected object was larger than or came next after the sample would simply require an intraverbal response to the instructional stimulus.

Reporting transformational responding. In many cases relations between stimuli are reported where there is no instructional stimulus, and the nature and name of the relation is left for the subject to determine. Tasks of this sort approach the requirements of problem solving in that it is left to the subject to find a transformational response that produces joint control between a sample stimulus and the product of the transformation. When one is found, it is frequently some aspect of the transformation that names the relation.

Thus, as illustrated in Figure 13, the phrase *they are all even* reports that all members of Series A enter into joint control with the intraverbal transformation *2-4-6,* whereas Series B would evoke the response *ordered* for the same reason with respect to the alphabetic intraverbal transformation. Series C, on the other hand, would evoke the autoclitic *disordered* in response to the conflict in topographies evoked by the letters and by all known transformations. Note also that in both cases the descriptions *ordered* and *disordered* are evoked by the serial nature of the joint control event; when there is a single joint control event, the terms *same, similar,* or *different* would be evoked.

Clearly, much remains to be explored here, but a beginning has been made, in that the mechanism here is reminiscent of the self-reporting in pigeons described by Lattal (1975). In the presence of a yellow keylight, pigeons could emit either a differential-reinforcement-of-low-rate (DRL) 10-s or a differential-reinforcement-of-other-behavior (DRO) 10-s response topography. After one or the other topography was completed, green and red keylights appeared, and selections of the color corresponding to the response topography just emitted (e.g., after DRL, select red; after DRO, select green) were reinforced. Presumably, just as the experimenter-controlled onset of red and green keylights terminated the DRL or DRO responding and evoked a report of the preceding behavior in pigeons, so may the subject-generated event corresponding to the onset of joint control evoke such a report in humans.

As broad categories, none of the preceding are mutually exclusive: There is no apparent reason why a subject might not emit two responses to the joint control event, and thus both describe a relation between stimuli and also act with respect to that relation. In doing so, in both describing what one is doing and doing it, one would normally be described as showing an awareness, consciousness, or understanding of one's own behavior. If this is so, then the preceding is not only a complete account of the behavior but also invites a fully operant account of these constructs.

Reports about the properties of a stimulus.[10] Besides describing relations between stimuli, joint control can evoke a category of autoclitics that describe the conditions under which a specified stimulus was found. Here again, the onset of joint control serves to mark the end of a behavior sequence whose features are subsequently reported. Thus, *it's hard* (or *easy*) *to find* reports that joint control only occurs after much (or little) scanning, whereas *it's hard to recognize* reports problems, but ultimate success, in achieving joint control between a self-duplic and a particular stimulus after it is encountered: Perhaps the speaker had trouble emitting tacts to the appropriate features of the stimulus.

A comment regarding the location of a sought-after object (*it's higher*) reports that

[10] The preceding section described autoclitics that reported on transformational responding in the self-duplic leg of joint control. This section describes autoclitics that report on events that occur in the tact leg of joint control.

joint control with the sample tact occurred after looking up. Closely related here are the autoclitics that report on the product of the scanning behavior rather than on the scanning behavior itself. Thus, *I have found it* reports that joint control has occurred (and thus the subject knows that this is the object sought). *I think this is it* reports about partial control. But what about *that's not it*? What is the stimulus when something is not present?

Reporting a Failure to Find

It has been pointed out elsewhere (Lowenkron & Colvin, 1992), that the inability of the unmediated selection account to describe responding in the absence of a specified stimulus is a serious deficiency. In this account, a response is made more probable in the presence of a particular stimulus as a result of its history of reinforcement in the presence of that stimulus. Other, incorrect, stimuli only function by *not* evoking the selection response, thereby allowing the correct stimulus to do so. But this does not fit with ordinary observation. People often respond to incorrect stimuli by doing something else. Thus, they may report that the stimulus is incorrect, they may look elsewhere, or they may change what they are looking for. How does this happen?

To account for such behavior only requires the assumption that a conflict between topographies that prevents joint control can act as a stimulus event just as well as the congruence of topographies that allows joint control. Occurrences of this conflict of control are as varied as the circumstances in which we respond to mismatches. Thus, we say a blue square is *not a red circle* because the description *red circle* does not enter into joint control with the tact for a blue square. One cannot look at a blue square and tact *red circle*.[11] In a similar fashion, the letter D is recognized as missing from the sequence A-B-C-E. The intraverbal recitation conflicts with the tact in the fourth position. This event, like joint control itself, may then evoke autoclitic responding: either verbal responses (*it isn't here* or *this isn't it*) or physical responses like pointing to stimuli or turning a page.

Oddity matching. The origins of this sort of behavior were explored by Lowenkron and Colvin (1992). Using overt, directly observable responses, the study showed that accurate, generalized matching to novel objects, not oriented in the same direction as the sample, depended on the absence of joint control between self-duplic rehearsals of the tact to the orientation of the sample and accurate tacts to the orientation of the correct comparison. Using a matching-task procedure similar to that illustrated in Figure 11, 4- and 5-year-old children learned to tact the orientation of the sample with the arrow, and then to maintain that orientation, unchanged, during comparison selection as selections of a comparison of dissimilar orientation were reinforced. When this behavior was acquired, the performance immediately generalized to symmetrical novel stimuli. Because selections with these stimuli had never been reinforced previously, accurate responding could not have resulted from any specific pairings of the arrow with these stimuli. It would thus seem that correct selections with these novel stimuli could be based only on the discrepancy between the tact (arrow orientation) they evoked as comparisons, and the tact they had evoked when they appeared as samples. Together with the earlier data, these findings indicate that it is possible to respond with a selection response to both the onset of joint control over a common topography (as illustrated by the generalization reported earlier) and also to the onset of a conflict between topographies.

Constant-relation matching. It has already been argued that operants that serve functions other than stimulus selection may be evoked by the presence or absence of joint control. As we shall see next, when these operants are transformational responses of the sort described earlier, there emerges a class of behavior with characteristics usually described as purposeful and goal oriented. Thus, as illustrated in Figure 12, Panel A (see also Lowenkron, 1989), after subjects were taught to select (under the instructional control of screen color) line lengths bearing various relations to the sample, they were exposed to a condition in which a line of the length specified by the transformation was not to be found. Under these conditions, subjects were trained to emit a second transformation, still under the control of the instructional stimulus, and to reselect.

When this behavior generalized to novel

[11] One might say *red circle*, but it would not be a tact.

stimuli, where selection was based on the width of rectangles or the distance between dots, the subjects, in effect, showed a generalized goal-oriented strategy. Thus, when instructed to find a rectangle or a dot distance one size larger than the sample, they sought it out. When they did not find it, they changed the basis of the search in an appropriate fashion and searched again, now seeking a comparison two sizes larger than the sample. In this way the subjects maintained the goal of finding a comparison larger than the sample, even with stimuli for which this was never trained.

A similar form of generalized goal-oriented responding was reported by Lowenkron and Colvin (1995) in the task described earlier requiring the selection of faces in response to spoken names (Figure 12, Panel B). Here too, the absence of a comparison specified by the current sample and the current instructional stimulus evoked appropriate transformational responding such that a new and appropriate target was sought based on intraverbal transformations using both the original as well as the new list (transfer set) of names. This is goal-oriented behavior at an exceedingly high level of abstraction!

Questions do remain here. Thus, in contrast to oddity matching, in which the conflict between topographies evokes selection, in constant-relation matching there is no discrete event that may be identified as the source of control over the transformational response. How this happens certainly warrants further exploration.

Goal-oriented behavior. The examples cited here appear to be goal oriented because failures to find a specified stimulus did not evoke the repetitive responding characteristic of extinction; rather, they immediately evoked a transformational response that modified the self-duplic and thus changed the stimulus being sought. The fact that subjects did this in a manner consistent with a previously specified relation (e.g., *larger than* or *next*), rather than in a random fashion, characterizes what would ordinarily be called a goal-oriented strategy intended to find a stimulus bearing a specified relation to the sample.

In general, then, at least some goal-oriented behavior may be said to result from the interaction of two autoclitics: Under joint control, subjects emit a selection response; when there is no joint control, they emit a transformational response that modifies the goal of the search. Some of the intelligence shown by a goal-oriented strategy thus resides in the effectiveness of the particular sequence of transformational responses the subject emits in the face of each failure to find the specified object. But behavior would show even more intelligence if a variety of transformations were available, each under the control of instances of partial success (i.e., instances of partial joint control between self-echoic rehearsals of the goal state and tacts describing the present state), because in this case the strategy appears to adapt appropriately to changing circumstances. Demonstrations of this sort are yet to be made, but it would certainly seem to involve nothing more than a straightforward extension of the demonstrated property of joint control to signal the nonoccurrence or the occurrence of some or all parts of a specified event.

Summary and Conclusion

It is possible to concisely summarize all of the performances discussed here because they all result from simple modifications of a common pattern. Figure 14, Panel A, illustrates the general case in which two stimuli (S1 and S2) evoke a common topography, and the autoclitic response serves to select a stimulus. If S1 and S2 are the same stimulus, then the performance illustrated is identity matching. If they share no physical features, but evoke a common response topography (Rv), then the figure illustrates bidirectional stimulus specification: S1 and S2 may each specify the other. If one stimulus is a word and the other is an object, then the performance illustrated is naming, or alternately, being named. If S1 contains more than one word and S2 is an object, then the words would be said to describe S2, and the figure illustrates recognition from a description, or bidirectionally, the selection of a description (e.g., a printed phrase) in response to an object. If the autoclitic response is not a selection response, it may be a report that the two stimuli do or do not match, or that the object has been recognized from its description.

Panel B illustrates cases in which sample stimuli are transformed before entering into joint control. When the autoclitic controlled by the joint control event is the selection re-

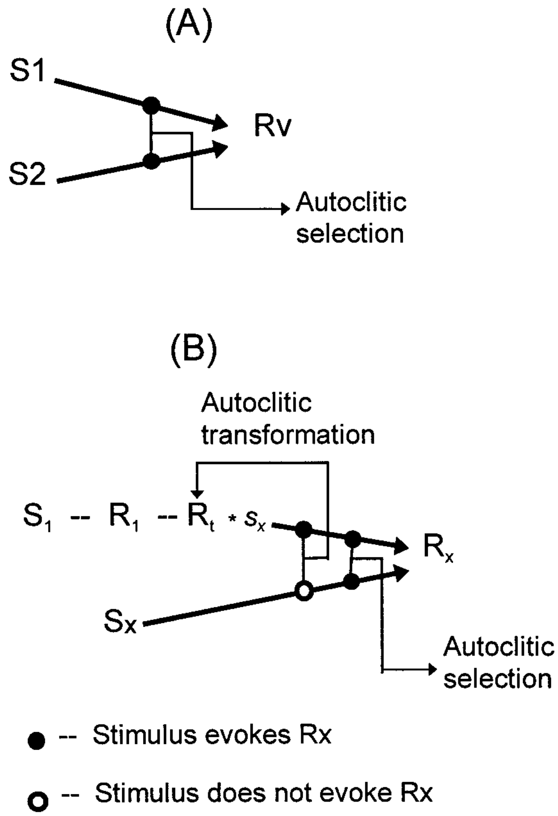

Fig. 14. Two models of joint control. Both describe a variety of different performances, illustrating here their structural similarities. (A) Simple joint control over a common verbal response (Rv) sets the occasion for a selection response. (B) Relational and goal-oriented matching. The initial response (R1) to a stimulus (S1) is modified by a transformational response (Rt) so as to produce s_x. In turn, s_x, by specifying the topography Rx, specifies the stimulus to be selected: Sx. Should no comparison be found that enters into joint control with s_x, this evokes Rt producing a different s_x, thus specifying a new stimulus (Sx). This goal-oriented loop continues until joint control occurs between some s_x and Sx, thereby occasioning the selection response.

sponse, the figure illustrates relational matching based on a transformation. If the autoclitic response is itself a transformational response, the figure illustrates goal-oriented behavior (looping repeatedly via the transformation, until determining, via the onset of joint control, that the goal has been reached, and terminating with the autoclitic selection response).

There is more. The reinforcing functions of joint control have not been discussed, although it seems certain that the correlation between the onset of joint control and reinforcement for correct selections can cause the joint control event to function as a conditioned reinforcer (Palmer, 1996). Likewise, the role of visual imagery has not been considered. Nonetheless, the foregoing suggests how a conceptual analysis might come to reclaim the ground behavior analysis has left, by default, to cognition.

It is perhaps appropriate to conclude by pointing out that unlike other recent contributions (e.g., Hayes, 1991; Horne & Lowe, 1996), this interpretive exercise is not the product of an ad hoc endeavor to provide a behavior-analytic account of some selected performances. Rather, the foregoing is an application of some concepts originally developed from empirical concerns that arose from efforts to directly train performances based on generalized abstractions such as identity, relative size, distance, orientation, negation, and order (Lowenkron, 1984, 1988, 1989; Lowenkron & Colvin, 1992, 1995). In the course of learning how to train these performances certain empirical conceptualizations regarding individual responses, sequences of responses, and stimulus control relations over these responses were found to be useful. The selection of a molecular analysis involving mediating responses, as opposed to a more molar analysis, was solely the result of these empirical concerns.

The extension of these concepts to the broad variety of performances described here was entirely unplanned; but over the course of experimentation (Lowenkron, 1984, 1988, 1989, 1991; Lowenkron & Colvin, 1992, 1995), as the concepts reappeared in various forms, they became increasingly better defined. With this clarity, their applicability to the interpretation of other behavior (as Skinner suggests in the opening quotation), became increasingly apparent. Likewise in the course of writing this paper did new vistas open; inviting the extension of this approach to problem-solving, rule-governed, and goal-oriented behavior. It is thus hard to believe that a concept so parsimonious in content, and yet so pervasive in explanatory function, will not play a significant role in a fully behavioral account of what is currently described as the role of language in human thought, but which we may ultimately come to understand as the role of verbal behavior in the behavior of the listener.

REFERENCES

Anderson, J. R. (1990). *Cognitive psychology and its implications* (3rd ed.). New York: Freeman.

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society, 106,* 135–163.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6,* 183–200.

Bucher, B. (1975). Establishing discriminative stimulus control within generalized oddity performance, without reinforcement. *Journal of Experimental Child Psychology, 20,* 38–50.

Cumming, W., Berryman, R., & Cohen, L. R. (1965). Acquisition and transfer of zero-delay matching. *Psychological Reports, 17,* 435–445.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behavior analysis in theory and practice: Contributions and controversies* (pp. 115–137). Hillsdale, NJ: Erlbaum.

Eikeseth, S., & Smith, T. (1992). The development of functional and equivalence classes in high-functioning autistic children: The role of naming. *Journal of the Experimental Analysis of Behavior, 58,* 123–133.

Ekstrand, B. R. (1966). Backward associations. *Psychological Bulletin, 65,* 50–64.

Goldstein, H., Angelo, D., & Whetherby, B. (1987). Effects of training method and word order on adults' acquisition of miniature linguistic systems. *The Psychological Record,* 89–107.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42,* 335–346.

Harnad, S. (1996). Experimental analysis of naming behavior cannot explain naming capacity. *Journal of the Experimental Analysis of Behavior, 65,* 262–264.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–46). Reno, NV: Context Press.

Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist, 47,* 1383–1395.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 185–241.

Horne, P. J., & Lowe, C. F. (1997). Toward a theory of verbal behavior. *Journal of the Experimental Analysis of Behavior, 68,* 271–296.

Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory and Cognition, 12,* 568–574.

Lattal, K. A. (1975). Reinforcement contingencies as discriminative stimuli. *Journal of the Experimental Analysis of Behavior, 23,* 241–246.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1954). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358–368.

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). Joint control and the generalization of selection-based verbal behavior. *The Analysis of Verbal Behavior, 9,* 121–126.

Lowenkron, B. (1996). Joint control and word-object bidirectionality. *Journal of the Experimental Analysis of Behavior, 65,* 252–255.

Lowenkron, B. (1997). The role of naming in the development of joint control. *Journal of the Experimental Analysis of Behavior, 68,* 244–247.

Lowenkron, B., & Colvin, V. (1992). Joint control and generalized nonidentity matching: Saying when something is *not. The Analysis of Verbal Behavior, 10,* 1–10.

Lowenkron, B., & Colvin, V. (1995). Generalized instructional control and the production of broadly applicable relational responding. *The Analysis of Verbal Behavior, 12,* 13–29.

McIntire, K. D., Cleary, J., & Thompson, T. (1989). Reply to Saunders and Hayes. *Journal of the Experimental Analysis of Behavior, 51,* 393–396.

Michael, J. (1982). Skinner's verbal operants: Some new categories. *VB News 1,* 1.

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

Michael, J. (1996). Separate repertoires or naming? *Journal of the Experimental Analysis of Behavior, 65,* 296–298.

Palmer, D. C. (1996). Achieving parity: The role of automatic reinforcement. *Journal of the Experimental Analysis of Behavior, 65,* 289–290.

Saunders, K. J. (1989). Naming in conditional discrimination and stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 51,* 379–384.

Schoenfeld, W. N., & Cumming, W. W. (1963). Behavior and perception. In S. Koch (Ed.), *Psychology: A study of a science* (pp. 213–252). New York: McGraw Hill.

Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behavior analysis in theory and practice: Contributions and controversies* (pp. 93–114). Hillsdale, NJ: Erlbaum.

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

Skinner, B. F. (Ed.). (1969). *Contingencies of reinforcement: A theoretical analysis.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1974). *About behaviorism.* New York: Knopf.

Urcuioli, P. J. (1985). On the role of differential sample behaviors in matching to sample. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 502–519.

Vygotsky, L. (1962). *Thought and language.* Cambridge: MIT Press.

Whitehurst, G. J. (1996). On the origins of misguided theories of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65,* 255–259.

Zentall, T. R., Edwards, C. A., Moore, B. S., & Hogan, D. E. (1981). Identity: The basis for both matching and oddity learning in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 70–96.

Received January 29, 1997
Final acceptance January 7, 1998